ACTUARIAL STATISTICS

COORDINATING EDITOR:

HARRY FREEMAN, M.A., F.I.A.

VOLUME II

CONSTRUCTION OF MORTALITY & OTHER TABLES

by

J. L. ANDERSON, B.A., F.I.A.

and

J. B. DOW, M.A., F.F.A.

CAMBRIDGE

Published for the Institute of Actuaries and the Faculty of Actuaries

AT THE UNIVERSITY PRESS

1952

PUBLISHED BY
THE SYNDICS OF THE CAMBRIDGE UNIVERSITY PRESS

London Office: Bentley House, N.W.1
American Branch: New York

Agents for Canada, India, and Pakistan: Macmillan

First Edition 1948
Second Edition 1952

Printed in Great Britain at the University Press, Cambridge
(Brooke Crutchley, University Printer)

CONTENTS

CHAPTER III

EXPOSED TO RISK FORMULAE

CHAPTER IV

EXPOSED TO RISK FORMULAE. AGE GROUPING BY DATE

CHAPTER V

EXPOSED TO RISK FORMULAE. AGE GROUPING BY EVENT

CHAPTER VI

PRACTICAL APPLICATION OF THE METHODS

CHAPTER VII

EXAMPLES OF THE CONSTRUCTION OF EXPOSED TO RISK FORMULAE

CHAPTER VIII

ERRORS OF APPROXIMATION

CHAPTER IX

THE CENTRAL RATE OF MORTALITY

CHAPTER X

THE CENSUS METHOD

CHAPTER XI

CONSTRUCTION OF SELECT TABLES

CHAPTER XII

USE OF CENTRAL RATES OF MORTALITY FOR SELECT TABLES. APPLICATION OF THE CENSUS METHOD

CHAPTER XIII

THE GENERAL THEORY OF SELECTION

CHAPTER XIV

THE CONSTRUCTION OF MORTALITY TABLES FROM LIFE OFFICE RECORDS

CHAPTER XV

HISTORY OF MORTALITY TABLES OF ASSURED LIVES

CHAPTER XVI

FORECASTING RATES OF MORTALITY. GENERAL THEORY OF PROJECTION

CHAPTER XVII

MORTALITY TABLES OF ANNUITANTS

CHAPTER XVIII

COMPARISON OF MORTALITY TABLES

CHAPTER XIX

DOUBLE- AND MULTIPLE-DECREMENT TABLES

CHAPTER XX

CONSTRUCTION OF NATIONAL LIFE TABLES

CHAPTER XXI

NATIONAL LIFE TABLES. SECTIONAL INVESTIGATIONS

CHAPTER XXII

SICKNESS RATES

INTRODUCTION TO VOL. II OF THE PART III TEXT-BOOK

THE issue of the present volume by Messrs J. L. Anderson, B.A., F.I.A., and J. B. Dow, M.A., F.F.A., completes the Text-Book for the Part III Examinations of the Institute and Faculty; the first volume, that on Statistics and Graduation by Mr Herbert Tetley, M.A., F.I.A., was published in 1946.

The services of all these authors, and of the co-ordinating editor, Mr H. Freeman, M.A., F.I.A., were acknowledged by our respective predecessors in office in their Introduction to Mr Tetley's volume; but we should like to add our tribute of gratitude for their unsparing labours, which have so greatly eased the task of present and future students for Part III. The present volume will, we believe, assist those students in obtaining a clear grasp of main principles, and will prove to be a valuable addition to actuarial literature.

A. H. ROWELL
President, The Institute of Actuaries
COLIN S. PENN
President, The Faculty of Actuaries

June 1947

EDITOR'S PREFACE

THE subject under the heading of the second volume of Actuarial Statistics is one which often causes difficulty to the student. One of the reasons for this is probably that, up to the present, the official reading has been diffuse. The need for an up-to-date text-book on the compilation and interpretation of mortality and similar tables has long been evident, and this book should satisfy that need.

Editing the book has been a most interesting task. Even when the authors of a text-book know their subject as well as do the authors of this one, they do not always express themselves logically and lucidly. In the present book the reader cannot fail to be struck with the clearness with which the successive arguments for the adoption of the formulae for the exposed to risk have been set out. The authors are at pains to stress that formulae are not to be learnt by rote but are to be evolved from first principles and devote a considerable portion of the book to the processes for solving the many different types of problem that beset the student in his first approach to the theory and practice of the construction of mortality tables.

H. F.

AUTHORS' PREFACE

THIS text-book has been written with special regard to the requirements of the student who is encountering for the first time the problems involved in the construction of mortality tables. We have therefore dealt at considerable length with the subject of Exposed to Risk Formulae which is a frequent cause of confusion at the first approach. The underlying theory has been built up from first principles and has then been applied to demonstrate the simplest type of formula. More complicated formulae have been developed by reference to the simpler types.

While some readers, with previous experience of the subject, may prefer a more general approach, we feel strongly that the average student should begin with the simplest cases and proceed by easy stages to the more complicated without any attempt to formulate a general method.

We have not thought it necessary to include any examples to be worked by students, but a number of representative examples have been included in the text to illustrate particular points.

We have had access to, and have made full use of, the notes provided for students by the Actuarial Tuition Service and we gratefully acknowledge our debt to their authors. We should also like to record our thanks to a number of our professional colleagues who have helped with advice and criticism, and to Mr H. Freeman who, as co-ordinating editor, has greatly eased the work of preparing the book for the press.

J. L. A.
J. B. D.

CHAPTER I

PRELIMINARY SURVEY

1. Introduction.

The mortality table is first introduced to the student of actuarial science as the instrument by which he can calculate probabilities of death and survivance and hence the values of monetary functions which involve these probabilities. It is put into his hands as a finished tool, and the theory and technique of its employment is the first part of his purely actuarial studies. The subject will be best known to many students under the title of "life contingencies". In all that follows, we accordingly assume that the reader has already studied this subject, and no attempt will be made to explain the nature of a life table or the meaning of the functions associated with it.

The apparently illogical course of asking a student to consider the uses of the mortality table before learning how or of what materials it is made is not without justification. In this way he can more readily appreciate the essential requirements of a satisfactory table. Moreover, he is able to complete the purely theoretical part of his studies before passing to a subject where practical considerations prevail. For, as we shall see, the construction of tables of mortality (or similar tables of sickness, marriage, withdrawal, etc.) is essentially a practical task involving as it does the choice, collection and analysis of statistical facts. It was the original, and is still one of the most important, tasks of the actuary. Major mortality investigations are rare events—rarer, perhaps, in this country than they should be—but investigations on a smaller scale must be made from time to time by the actuary who would base his conclusions on something more solid than surmise.

Why, it may pertinently be asked at this stage, must the construction of mortality tables and the calculation of rates of mortality be treated as problems in practical statistics? Is it not possible to arrive by theoretical reasoning at some universal law of mortality, some comprehensive formula for the rate of mortality which could be adapted to any body of lives by assigning suitable values to the

constants? The answer is that although in some investigations the progression of the rates of mortality from age to age has been found to conform fairly closely to one mathematical formula or another—as, for example, those of Gompertz and Makeham—no law has yet been discovered which is universally applicable. Whether such a law exists is an open question, but when we consider how many factors are commonly known or suspected to influence mortality—age, sex, occupation, heredity, year of birth, place of residence, to name but a few—it is hardly surprising that the theoretical method of approach has met with such little success and that we are compelled to adopt the statistical method of recording the results of certain definite investigations and drawing from them such general conclusions as appear reasonable and justified.

2. Objects of a mortality investigation.

The first and most obvious purpose of a mortality investigation is to ascertain and put on record the rates of mortality experienced by a particular body of lives over a particular period. But we must also consider what use we may wish to make of these records once they have been set up. This problem requires careful thought.

Whatever opinion we hold regarding the existence of a law of mortality, we probably admit to an intuitive belief that there is some measure of order underlying the rates of mortality that we study. We expect, for example, that in any normal experience the rates of mortality will be higher at the older ages than at the younger. We may be prepared to go further and agree that the information we obtain from investigating the mortality of one group of lives will assist us in estimating the mortality of another group apparently subject to similar influences. This intuitive belief is fundamental in nearly all our work on mortality tables. Most office premiums are calculated with the help of mortality tables based on the experience of lives assured in the past. The financial provisions of much legislation affecting social services, such as National Insurance, are based on mortality and other tables derived from the results of past censuses. In both cases, the assumption made is that the future mortality of certain more or less clearly defined groups—such as lives accepted by assurance companies at tabular rates, old age pensioners, etc.—will be similar to or will differ in

some systematic and predictable way from the mortality experienced by similar groups in the past.

If this assumption is accepted as a not unreasonable hypothesis, the purpose of our investigations is immediately widened. The mortality table ceases to be merely a statement in convenient statistical form of certain observed facts and becomes also a means whereby we can estimate or predict the mortality of lives who have not been directly observed. Our aim, in fact, is not only to record but also to compare, where by comparison we mean any process which sets up the mortality of one group of lives as a measure of, or in contrast to, the mortality of another group. This includes uses of the mortality table already referred to; it also embraces attempts to find a general law of mortality (in which we set the results of our observations against the values obtained by a mathematical formula) and investigations into the mortality of groups of lives whom we believe to be in some respect dissimilar. The latter category includes the important type of investigation where we compare the mortality of assured lives revealing a particular feature in their medical or family history with the mortality of a standard group with no such impairment. We shall see that this extension of our aims greatly increases the difficulty, interest and value of our task.

3. Nature of the data required.

In its columns of l_x and d_x, the mortality table has been said to record the "march of a generation through time". (The phrase was originally used by Dr Farr, whose work in the construction of mortality tables is dealt with in Chapter XX.) It suggests that a large group of people born at the same moment has been under observation and that the numbers surviving to each birthday have been recorded. If no lives have been added to the group and none have left it, except by death, the numbers recorded will form, without adjustment, the l_x column of the mortality table; and the differences between them will give us the numbers dying at each age, i.e. the d_x column. Unfortunately, the making of such observations presents practical difficulties. In the first place, if our mortality table is to be complete, we must extend our observations to cover the whole lifetime of all members of the group; secondly,

it is difficult, if not impossible, to arrange that, throughout the period of investigation, there will be no migration to or from our group.

We are, therefore, driven to seek some less direct line of attack, and our attention naturally turns to the functions expressing the probabilities of death and survivance, q_x, p_x, m_x, etc. Clearly, if one of these functions is known at each age, we can readily construct the l_x and d_x columns. Given q_x, for example, the relations

$$d_x = q_x l_x \quad \text{and} \quad l_{x+1} = l_x - d_x$$

provide us at once with what we require. Now, if we observe for a period a group containing lives of all ages and record the numbers living and the deaths at each age, we can find the rates of mortality of the group and hence our l_x and d_x columns. This method, which avoids the practical difficulties mentioned above, has been so generally used that it has come to be regarded as the normal method. It is well, however, to note that it has been chosen for practical and not for theoretical reasons, and that its adoption involves an important change in our point of view.

In the first method, we traverse the years with a single group of lives so that those under review at age $x+n$ are the actual survivors of lives observed n years previously at age x. In the second method this is no longer true, since we confine our observations to a comparatively short period of time and observe lives of all ages simultaneously. This raises some important points, both theoretical and practical, which will be discussed in later chapters, but for the present we accept the second method as the one in practical favour and ask only that we shall be supplied with the numbers living at each age over a short period and the number of deaths among them at each age during the same period.

4. Sources of the data.

The necessary information for a mortality investigation can be obtained from one of two main sources, which we may broadly class as public and private. In the former category we include the information given in census returns or any enumeration of population in a country, district or town and the relative records of deaths; in the latter are included all the records of assurance companies, friendly societies, professional bodies, superannuation

funds, etc. Both sources give us what we need for our purpose, but in entirely different forms. From the public records, we generally obtain information regarding lives in the mass. Census figures, for example, tell us the numbers living at each age or in each age group on a specified date, but no information will be available regarding individual lives. On the other hand, private records will usually give particulars such as date of birth and death for each person included, and we shall be able, if we wish, to deal with each life separately in a way which would be impossible with census or similar data. This, we shall see, leads us to devise different methods of approach in the two cases.

5. Choice of data.

Our discussion of the nature of the data required left unanswered the fundamental question, "How can we best determine the limits of the group to be observed?" To this question no simple and final answer can be given. In many cases our group is determined for us, as, for example, in dealing with census figures, where the group is the whole population: in others the composition of our group will depend on the objects of our investigation. If, as will frequently be the case, we wish to use our results as a guide to the mortality of another group of lives, instinct and theory alike counsel us to ensure, so far as we can, that the two groups resemble each other in all essential characteristics. Thus we would not use a mortality table based on a census of the whole population to estimate the mortality of a group of lives assured, since, if there is any virtue in the process by which offices select the lives they assure, the two groups differ fundamentally, and there is no reason why the mortality of one should resemble closely the mortality of the other. Neither would we use the results of an investigation into the mortality of a group of lives resident in one country as a guide to the mortality of lives who reside in another, because the two groups differ in one factor which may exercise an important influence on mortality.

If, on the other hand, we wish to compare the mortality of lives in two different countries, we require to find two groups of lives, one in each country, as nearly as possible similar in constitution as regards sex, occupation, etc., so that any differences in mortality

revealed by the comparison may reasonably be attributed to differences in climate and environment.

6. Homogeneity of data.

In general, the more clearly defined and uniform the characteristics of the lives included in a group, the greater will be the value of the results achieved by the investigation. A group in which all the lives resemble each other in all characteristics affecting mortality is said to be *homogeneous*. Such a group is a theoretical ideal which cannot be fully realized in practice. What we can and must do is to ensure that all major causes of heterogeneity are excluded and that the resulting group is subject to no special influences which would reduce its value as a measure of the mortality of other groups of similar type. Clearly this is a subject on which no definite rules can be laid down and where the actuary must be guided by his individual judgement and experience.

7. Errors due to insufficient data.

Our quest for homogeneity may reduce substantially the size of our groups and, if we carry the process too far, our results may be vitiated by the random errors which arise from dealing with too scanty data. To deal with this subject fully would involve a wide discussion of statistical theory which would be out of place here. The point is so important, however, that a brief explanation is desirable if only for the sake of completeness.

If we had a large body of lives forming as nearly as possible a homogeneous group (as defined in the preceding paragraph), and if we divided this into small groups with a hundred lives in each, we should not expect that, over a period of a year, the same number of deaths would occur in every group. Any doubt on this point can be set at rest by the *reductio ad absurdum* argument of taking groups with one life in each, in which case the same number of deaths in each group would necessitate either none or all of the lives dying within the year. The theory of statistics shows that, as the size of the groups is increased, closer agreement in the number of deaths in each group becomes likely, and similarly that, if the proportion of deaths in a group is compared with the proportion in the whole body, then, in general, the larger the group the closer will be the agreement.

When investigating rates of mortality, therefore, we must remember that different rates might have resulted if a larger body of data had been available. By the application of statistical theory, the process of graduation enables us to approximate to the rates which would have been obtained in ideal circumstances, i.e. if an infinitely large number of lives of the particular type under review had been available. The subject of graduation is outside the scope of this volume and we shall not attempt to discuss it here. It is mentioned only so that the reader will realize that an attempt can be made to overcome the disadvantages arising from paucity of data and to reassure him if he feels that there is no way of bridging the gap between the ideal conceptions of theory and the results of actual experience.

8. Summary of the process.

The various stages in the process of constructing a mortality table will now be summarized before they are discussed in detail.

(*a*) The scope of the investigation must be decided, i.e. the lives to be included, the period to be covered and the subdivisions of data to be made. This in turn depends on the object of the investigation, which may determine the scope for us (as, for example, when we wish to investigate the mortality of the whole population of Great Britain over a given period) or may raise many controversial questions (as in a life office investigation). In the latter event, the problem of subdividing the data is a notoriously difficult one, since, as we have seen above, two aims which in practice are mutually inconsistent must be kept in view, i.e. the inclusion of as large a body of data as possible and the attainment of perfect homogeneity. In practice, a compromise is sought by subdividing the data until each subgroup is reasonably homogeneous and investigating its mortality separately. Unfortunately, the numbers in individual subgroups may then be too small to justify any reliance on the results, and it is no easy matter for the actuary to achieve a suitable balance between these opposing forces.

(*b*) The data must be collected, analysed and tabulated in a form suitable for calculating what are known as the *crude rates of mortality*, i.e. the rates obtained direct from the data without adjustment.

(*c*) The crude rates of mortality must be graduated. The reasons for and the aim of this process have been indicated above. Once the graduation is complete, the mortality table is ready for use, either for the calculation of some of the many functions involving probabilities of life and death, or of monetary functions such as premiums.

9. Conclusion.

Enough has probably been said to show that the subject involves problems of a varied nature and that it is not merely a logical extension of mathematical theory. A firm grasp of the underlying theories is, of course, essential, but it must never be forgotten that the subject is a practical one. Many of the problems encountered are of a controversial nature, and the possession of a keen critical faculty is therefore valuable. Moreover, as the reader will have gathered already, the actuary engaged on the construction of a mortality table will constantly be reduced to making approximations, partly because his data may be deficient and partly to reduce his work to reasonable limits. The small amount of mathematics involved is rarely employed except as a device to cut down arithmetical work. In short, the construction of a mortality table involves an attempt to fashion a practical instrument in the face of numerous and sometimes irreconcilable considerations. The actuary engaged in the task can best hope to achieve his object by maintaining a clear view of the underlying principles, by the exercise of great care and by the application of that uncommon quality paradoxically called common sense.

Finally, the reader should not be dismayed if he has found some of the conceptions in this chapter difficult to understand. The summary is intended mainly for those who like to have a bird's-eye view before traversing the landscape on foot, and the greater part of what has been said will be explained in more detail in the succeeding chapters.

CHAPTER II

CALCULATION OF THE RATE OF MORTALITY FROM FIRST PRINCIPLES

1. If we were to study the construction of mortality tables in logical order, the first stage would be to consider the choice, collection and analysis of the data. It is difficult, however, to discuss these aspects of the problem satisfactorily without a clearer knowledge of the form in which the analysed data are required. For the present, therefore, we shall assume that the exact scope of the experience has been decided and that our immediate problem is the actual calculation of the mortality functions. The word "experience" is used here in a special sense. It denotes an investigation of the mortality, sickness, etc. experienced by a given body of lives over a given period, and in this sense it will be encountered frequently in actuarial literature.

We shall also assume that any subdivisions of the data necessary to achieve a reasonable degree of homogeneity have been made, so that we are faced with the problem of measuring, for the lives in a particular subdivision, the annual rate of mortality, q_x, at each age for which there are data. q_x is the only function which we need consider, since the method of calculating the central rate of mortality, m_x, differs only in detail while, on the other hand, the force of mortality, μ_x, is difficult to calculate by practical methods and is seldom used.

2. Definition of the rate of mortality.

It may be well at this point to recall exactly what is meant by the annual rate of mortality—more often called simply the rate of mortality. This function may be defined in various ways, but for the present we shall adopt the following:

If θ_x deaths occur between exact ages x and $x+1$ among a group of P_x lives of exact age x who remain under observation until the attainment of exact age $x+1$ or who would have so remained if they had

not passed out of observation by death previously, then the annual rate of mortality experienced by these lives at exact age x is $q_x = \frac{\theta_x}{P_x}$.

The condition that all the lives remain under observation until age $x+1$ is of particular importance. If some of the lives were observed only until age $x+t$ (where $t<1$), we could not tell whether any deaths among them occurred between ages $x+t$ and $x+1$. This condition expresses a fundamental principle, namely that we must not include in P_x any lives among whom deaths may have occurred without our knowledge, nor in θ_x deaths occurring among lives who were not included in P_x. The deaths included in θ_x and the lives included in P_x must correspond exactly. This principle appears simple and indeed obvious, but its neglect is a frequent source of error even among experienced operators.

3. Application of the definition.

The phrase "P_x lives of exact age x" which occurs in the definition of q_x naturally calls to mind a group of lives attaining age x on the same day, so that the succeeding year of life is the same as the succeeding year of time. Actually the definition does not require this restriction. We can include in our group lives who attain age x at different times so long as they all remain under observation for the whole year of life from age x to age $x+1$. This allows us to extend our observations over as long a period as we please and to include all lives who within that period pass from age x to age $x+1$. An example may make this clearer.

Suppose that the investigation covers the calendar years 1930 to 1934 inclusive.

Count all the lives attaining age x between 1st January 1930 and 31st December 1933 inclusive and denote this number by P'_x. Those attaining age x in 1934 are excluded, because they do not attain age $x+1$ within the period of the investigation and the definition insists that the lives included in P'_x can be traced until the attainment of age $x+1$.

Count the number of deaths between exact ages x and $x+1$ occurring among the lives included in P'_x and denote this by θ'_x.

Then $$q_x = \frac{\theta'_x}{P'_x}.$$

It will be observed that θ'_x does not include all deaths between ages x and $x+1$ in the five-year period, but excludes

(*a*) deaths among lives attaining age x in 1929; and

(*b*) deaths among lives attaining age x in 1934.

The lives among whom these deaths occurred were not included in P'_x and the deaths must, therefore, be excluded from θ'_x.

In this example, q_x is the rate of mortality experienced at age x by the lives who attained that age in the calendar years 1930 to 1933, all these lives being assumed to remain under observation until age $x+1$ or previous death. We must not fail to notice, however, that the example involves the assumption that lives attaining age x on 1st January 1930 are subject to the same rate of mortality as lives attaining age x on 31st December 1933. If this assumption is untrue, the data are not homogeneous and further subdivision is required. When the period of observation is short, this point has little practical importance, but when investigations cover, as they have done in the past, periods as long as thirty years, the mortality of lives attaining a given age at the beginning of the period may differ widely from that experienced by lives of the same age at the end. Failure to take account of this difference may lead to wrong conclusions being drawn from the results.

4. Fractional periods of exposure.

The definition of q_x requires that all lives must be under observation from age x to age $x+1$ or previous death. It is interesting to see what effect this requirement has had on the investigations of the previous paragraph. It has led us to exclude:

(1) all lives who attained age x in 1929; and

(2) all lives who attained age x in 1934.

It is obviously desirable to include these lives, if we can do so without having to resort to unjustifiable approximations, since we thereby increase the data by about 25 per cent and thus improve the statistical value of our results. All the lives were under observation for part of the year of age x to $x+1$, the former from 1st January

1930 to age $x+1$ and the latter from age x to 31st December 1934, and we have records of all the deaths which occurred among them during these fractional periods.

Consider two groups of lives, one consisting of lives who attained age x a fraction of a year s before 1st January 1930 ($0<s<1$) and the other of lives who attained age x a fraction of a year t before 1st January 1935 ($0<t<1$).

The obvious way of incorporating the former group is to include a fraction $1-s$ in the denominator for each life and to include in the numerator the number of deaths in the group between ages $x+s$ and $x+1$, i.e. in a fractional period of duration $1-s$ years. Similarly, for the second group, t would be added to the denominator for each life and the number of deaths between ages x and $x+t$ added to the numerator. It should be noted that the numerator would include all the deaths between ages x and $x+1$ within the period of the investigation, when the groups applicable to all values of s and t had been incorporated in this way.

The procedure suggested above may be more easily understood if the two expressions for q_x are set out by means of symbols.

Let P_x^{29}, P_x^{30}, etc. be the number of lives who attained age x in 1929, 1930, etc., and θ_x^{29}, θ_x^{30}, etc. the number of deaths among them between x and $x+1$. (It must be stressed that θ_x^{29}, θ_x^{30}, etc. do not represent the number of deaths in the years 1929, 1930, etc.)

Then the first value we obtain for q_x (para. 3) is given by

$$q_x = \frac{\theta_x^{30}+\theta_x^{31}+\theta_x^{32}+\theta_x^{33}}{P_x^{30}+P_x^{31}+P_x^{32}+P_x^{33}} = \frac{\theta'_x}{P'_x}. \qquad \ldots\ldots(1)$$

Now, let ${}^sP_x^{29}$ be the number of lives attaining age x in 1929 with birthdays a fraction s of a year before 1st January 1930 and ${}^{1-s}\theta_x^{29}$ the number of deaths among them *on or after* 1st January 1930 and before the attainment of age $x+1$.

Then by the argument of the preceding paragraph we include ${}^{1-s}\theta_x^{29}$ in the numerator and $(1-s)\,{}^sP_x^{29}$ in the denominator of the fraction for q_x.

Similarly, let ${}^tP_x^{34}$ be the number of lives attaining age x in 1934 with birthdays a fraction t of a year before 1st January 1935 and ${}^t\theta_x^{34}$ the number of deaths among them after the attainment of

age x and before 1st January 1935. For these lives, ${}^t\theta_x^{34}$ will be included in the numerator and $t\,{}^tP_x^{34}$ in the denominator.

This applies for all values of s and t between 0 and 1, and hence we have the following formula for q_x:

$$q_x = \frac{\sum_{s=0}^{1} {}^{1-s}\theta_x^{29} + \theta_x^{30} + \theta_x^{31} + \theta_x^{32} + \theta_x^{33} + \sum_{t=0}^{1} {}^t\theta_x^{34}}{\sum_{s=0}^{1} (1-s)\,{}^sP_x^{29} + P_x^{30} + P_x^{31} + P_x^{32} + P_x^{33} + \sum_{t=0}^{1} t\,{}^tP_x^{34}}$$

$$= \frac{\theta_x}{\sum_{s=0}^{1} (1-s)\,{}^sP_x^{29} + P'_x + \sum_{t=0}^{1} t\,{}^tP_x^{34}}, \qquad \ldots\ldots(2)$$

where θ_x is the number of deaths between ages x and $x+1$ occurring between 1st January 1930 and 31st December 1934 and the summations extend over all groups for which data are available.

5. Assumption underlying the method of dealing with fractional exposures.

Before we leave formula (2), we must realize that we have made an important assumption. Consider a group of ${}^tP_x^{34}$ lives in respect of whom we have included in the denominator $t\,{}^tP_x^{34}$, i.e. a fraction t for each life under observation from age x to age $x+t$ or previous death, and in the numerator ${}^t\theta_x^{34}$, i.e. the number of deaths among these lives during the fraction t of a year from age x to age $x+t$. This procedure would be justified if we knew that, given an adequate amount of data, the deaths among a group of lives under observation from age x to age $x+1$ or previous death would be uniformly distributed over that year of age, i.e. that the deaths between ages x and $x+t$ would be in the ratio $t:1$ to the total deaths between ages x and $x+1$. This is a subject which requires further investigation and we shall deal with it more fully at a later stage. In the meantime, the reader is asked to accept the assumption as being in general sufficiently correct for practical purposes.

6. Exposed to risk at age x.

The inclusion of fractions of a year of life in the denominator of the ratio which gives q_x introduces a new conception, i.e. the number of years of life as opposed to the number of lives. The

denominator now gives the number of years of life between x and $x+1$ during which the lives attaining age x in 1929–34 inclusive were in the aggregate exposed to the risk of death in the years 1930–34, each death between x and $x+1$ being included to the same extent as if death had not occurred. This is called *the exposed to risk at age x* and is generally denoted by E_x. The method of dealing with deaths when calculating E_x is important and should be noted carefully. In particular, it should be noted that a fraction $1-s$ is included for a life who would have been aged $x+s$ on 1st January 1930, if death had not occurred in 1929 after the birthday in that year.

It will be seen that, whereas the denominator of q_x has previously consisted of the number of lives attaining exact age x, E_x is the number of years of exposure between exact ages x and $x+1$, i.e. the one consists of the number passing through a particular point of age and the other of the number of years of life between two successive ages, deaths being given the same exposure as if they had survived. The two conceptions are, however, consistent for, in the special case where the lives are all under observation for the full year of life, the number of lives is the same as the number of years of exposure.

No change in the definition of q_x is either necessary or desirable. E_x should rather be regarded as an approximation to the number of lives of exact age x who, if observed for a year, would have given rise to the recorded number of deaths between exact ages x and $x+1$.

7. Approximations to lengths of fractional periods of exposure.

The calculation of E_x would be a laborious process if the functions $\sum\limits_{s=0}^{1}(1-s)\,{}^{s}P_x^{29}$ and $\sum\limits_{t=0}^{1} t\,{}^{t}P_x^{34}$ were to be evaluated accurately. It is therefore usual to approximate to the average values of s and t for all the lives included in P_x^{29} and P_x^{34}. One way of doing this is to take random samples of the lives—for example a proportion of perhaps 5 or 10 per cent chosen from P_x^{29} and P_x^{34} at random—and to calculate the average values of s and t from the samples. These values would then be taken as approximations to

the corresponding values for the whole of P_x^{29} and P_x^{34}. E_x then becomes $(1-s)\,P_x^{29}+P'_x+tP_x^{34}$.

The values of s and t depend on the distribution of the birthdays of the lives in P_x^{29} and P_x^{34}. Although these two groups are mutually exclusive, they would usually consist of lives of the same type and much the same distribution of birthdays over the calendar year would be expected, so that the same value could probably be assigned to s and t. In practice, it is customary to assume that $s=t=\frac{1}{2}$ (i.e. that the birthdays of both groups of lives are uniformly spread over the calendar year) and it is obvious from general considerations that this will not be far from the truth if a fairly large number of lives is involved. For the sake of greater generality, we shall avoid this assumption in what follows and adopt the form

$$E_x=(1-s)\,P_x^{29}+P'_x+sP_x^{34},$$

which reduces to the more symmetrical expression

$$E_x=(1-s)\,P_x^{29}+P_x-(1-s)\,P_x^{34} \qquad \text{......(3)}$$

if we substitute P_x for $P'_x+P_x^{34}$.

8. Alternative form of formula (3).

Exception may be taken to the inclusion of the function P_x^{29} on the ground that 1929 is outside the period of the investigation. Both for this reason and for reasons of practical convenience it is customary to substitute for P_x^{29} and P_x^{34} the numbers of lives aged between x and $x+1$ on 1st January 1930 and 31st December 1934 respectively. These are known as the lives existing at the beginning and end of the investigation or, more simply, as the *beginners* and *enders*. If we denote the numbers by b_x and e_x, formula (3) becomes

$$E_x=(1-s)\,b_x+P_x-(1-s)\,e_x. \qquad \text{......(4)}$$

We shall discuss in Chapter VIII the nature of the assumption involved; in the meantime it is sufficient to note that the error introduced is unlikely to be large, especially as the separate errors caused by substituting b_x for P_x^{29} and e_x for P_x^{34} act in opposite directions.

9. Withdrawals and new entrants.

Up to the present, it has been assumed that all lives remain under observation from age x to age $x+1$ or until previous death, except in so far as any part of this year of age falls outside the period of the investigation. In practice, lives will come under and pass out of observation for various causes and seldom at integral ages. If, for example, the lives included in the investigation are those assured with a life office, the number under review will be increased by the granting of new policies and reduced by the surrender of existing policies and the maturity of endowment assurances, and these events will not usually occur at integral ages.

Lives passing out of observation for any reason other than death or survival to the end of the investigation period are usually called *withdrawals*. The withdrawals between ages x and $x+1$ are similar to the lives aged between x and $x+1$ at the closing date of the investigation in that each of them was under observation between ages x and $x+1$ for a fraction of a year of age from age x to the age at exit. The same assumption as before about the distribution of the deaths over the year of age therefore leads to the inclusion of these fractions in E_x and, if w_x be the number of withdrawals between ages x and $x+1$ and the average age at withdrawal be $x+h$, the total exposure to be included in E_x for the withdrawals will be hw_x. Formula (4) in para. 8 includes a unit for each of these lives and the term $(1-h)\,w_x$ must, therefore, be deducted to reduce the exposure to hw_x. This argument requires special consideration in the case of withdrawals before the birthday in 1930 or after the birthday in 1934 (see VIII, 7).

In the same way, the lives entering during the period of the investigation, usually known as *new entrants*, are akin to the lives who were under observation at the commencing date. The exposure at age x is, therefore, taken as $(1-k)\,n_x$, where n_x is the number of new entrants between ages x and $x+1$ and the average age at entry is $x+k$. Formula (4) of para. 8 does not contain any allowance for the new entrants between ages x and $x+1$, as none of them either attained age x while under observation or was under observation on 1st January 1930. The term $(1-k)\,n_x$ must therefore be added, special consideration being again necessary for

entrants before the birthday in 1930 or after the birthday in 1934. The adjusted form of the expression, when there are both new entrants and withdrawals, is accordingly:

$$(1-s)\,b_x+P_x+(1-k)\,n_x-(1-h)\,w_x-(1-s)\,e_x. \quad \ldots\ldots(5)$$

10. Errors of approximation.

It was stated in para. 6 that each death is included in the denominator to the same extent as if death had not occurred. In formulae (4) and (5) we have given to all deaths between ages x and $x+1$ in the period of the investigation full exposure from age x, or such later age at which the life came under review, to age $x+1$. This is not always the same as the exposure according to the previous rule, although the practical effect of the difference is usually negligible. We shall defer consideration of the problem until a later chapter (VIII, 4).

The reader may perhaps feel that he is being asked to take too much for granted and that the outstanding points should be cleared up forthwith in order that he may be satisfied that the formulae which have been devised are sufficiently accurate. To do this would, however, seriously interrupt the development of the main argument. A full investigation of the points in question is necessarily complicated and it is undesirable to go into great detail at such an early stage, when it is essential to concentrate on the main principles.

11. Alternative definition of q_x.

The annual rate of mortality may be defined in various ways and, although the definition in para. 2 is as simple as possible and therefore makes a suitable starting-point, some readers may prefer the following as being in a more conventional form:

The annual rate of mortality experienced by a group of lives at exact age x is the ratio of the number of the deaths occurring among these lives between exact ages x and $x+1$ to the number of lives attaining exact age x, there being no causes of decrement other than death and no causes of increment.

As in the earlier definition, no attempt has been made to define q_x in such a way that the method of dealing with fractional periods will

follow directly from the definition. Any such definition would have to depart from the fundamental conception of the rate of mortality. The exposed to risk when there are fractional exposures should therefore be regarded as an approximation on the lines indicated in the last sentence of para. 6.

The reader should have no difficulty in satisfying himself that the two definitions have the same meaning. The earlier definition is, in fact, the expression in symbols of the definition now given.

12. The life table.

Once values of q_x have been obtained over the available range of ages, the process of graduation is carried out as explained in Chapter I. Thereafter, a *life table* can be built up by starting with any suitable radix and applying successively the relations

$$d_x = l_x \times q_x, \quad l_{x+1} = l_x - d_x.$$

CHAPTER III

EXPOSED TO RISK FORMULAE

1. Connection between exposed to risk at successive ages.

In the last chapter we evolved a method for obtaining the values of E_x and θ_x and hence of q_x for each age x. We dealt with each age separately and found E_x and θ_x directly from the original data. It is clear, however, that if the investigation covers a period of more than one year, the same life may contribute to the values of E_x at several successive ages. In an experience covering n years, for example, a life aged x in the first year of the experience may be included in the exposed to risk at each age from x to $x+n-1$. This suggests that some convenient relationship may exist between the values of E_x at successive ages, which will enable us, when we have obtained the value for one age, to derive other values by a continuous process, thus reducing the work.

2. Exposed to risk when there are no fractional periods of exposure.

Let us begin by considering the simple example discussed in (II, 3), in which lives come under observation on their birthdays in 1930 and remain under observation until their birthdays in 1934 or previous death, and let us suppose that, for a particular value of x, E_{x-1} has already been obtained by counting the number of lives who attained age $x-1$ in the years 1930–33 inclusive. If we were to set out to obtain E_x by enumeration, we should find that the same lives would be included as in our enumeration of E_{x-1} with the following exceptions:

(*a*) Those who attained age $x-1$ in 1929 and who survived to age x in 1930 must be added, since lives who attained age x in 1930 are included in E_x.

(*b*) Those who attained age $x-1$ in 1933 and who survived to age x in 1934 must be excluded, since E_x must include only those who attained age x in the years 1930–33 inclusive.

(*c*) Those dying at age $x-1$ last birthday between their birthdays in 1930 and 1934 must be excluded; they are treated as exposed for a full year at age $x-1$, but are not exposed at all at age x.

By making these adjustments we can transform the group of lives constituting E_{x-1} into the group constituting E_x. Algebraically we have

$$E_x = E_{x-1} + b_x - e_x - \theta_{x-1}, \qquad \ldots\ldots(1)$$

where b_x = the number of lives who attained age x in 1930 (see (*a*) above),

e_x = the number of lives who attained age x in 1934 (see (*b*) above),

and θ_x has the meaning assigned to θ'_x in (II, 3).

3. Underlying theory of exposed to risk formulae.

A relation of the type of (1) above is known as an *exposed to risk formula*. Such formulae are of the greatest practical importance. If the value of E_x is found for the youngest age included in the experience, then values of E_x for all subsequent ages may be obtained by the comparatively simple process of calculating at each age values of b_x, e_x and θ_x without direct reference to the numbers who attain the various ages in the calendar years 1931–33 inclusive.

The theory underlying the formula is simple. It is no more than the algebraic expression of the obvious fact that a life coming under review at age x (i.e. as a member of b_x) remains under observation and is therefore included in E_x, E_{x+1}, ... until death or until the end of the period covered by the experience (i.e. until included in one or other of the groups θ_x, θ_{x+1}, ... or e_{x+1}, e_{x+2}, ...). Each life is included once among the beginners and once among either the deaths or the enders; i.e. over the whole range of ages included in the experience the total number of entrants must equal the total number of exits and hence, when we reach the age after which no movements occur, E_x must be zero. If this is not so, there is an error in the work. Notice also the relationship

$$E_x = \sum^{x} b_y - \sum^{x} e_y - \sum^{x-1} \theta_y, \qquad \ldots\ldots(2)$$

where $\sum^{x}$ denotes the summation over all ages from the youngest

age for which data are available up to and including age x. At the youngest age (α say), the exposed to risk formula reduces to the relation $E_\alpha = b_\alpha$, which could of course be obtained direct from first principles, since obviously e_α and $\theta_{\alpha-1}$ must be zero or α would not be the youngest age in the experience. It should be noted that the word "movements" is used above in the sense of movements relative to the experience, i.e. lives coming under or passing out of observation.

4. Numerical example.

Suppose that our investigation comprises the following lives and that we wish to calculate E_x for all ages for which data are available.

Life	Date of birth	Date of death (if in 1930–34)	Age attained in 1930	Age attained in 1934
A	1st Mar. 1900	—	30	34
B	3rd July 1898	17th June 1931	32	—
C	25th May 1898	—	32	36
D	19th Dec. 1900	—	30	34
E	13th Nov. 1899	12th Dec. 1934	31	35
F	2nd Sept. 1899	—	31	35
G	15th Feb. 1896	—	34	38
H	1st Aug. 1900	21st May 1932	30	—

Remembering that all lives enter the experience on their birthdays in 1930 and remain under observation until their birthdays in 1934 or previous death, we may schedule the lives thus:

Age x	b_x Includes	b_x No.	e_x Includes	e_x No.	θ_x Includes	θ_x No.
30	A, D, H	3	—	—	—	—
31	E, F	2	—	—	H	1
32	B, C	2	—	—	B	1
33	—	—	—	—	—	—
34	G	1	A, D	2	—	—
35	—	—	E, F	2	—	—
36	—	—	C	1	—	—
37	—	—	—	—	—	—
38	—	—	G	1	—	—

E died after his birthday in 1934 and his death is therefore outside the experience. B died between the ages of 32 and 33 and is therefore included in θ_{32}. H died between the ages of 31 and 32 and is therefore included in θ_{31}. The calculation of E_x will then proceed thus:

Age x	b_x (1)	e_x (2)	θ_{x-1} (3)	(1)–(2)–(3) (4)	$E_x=E_{x-1}$ +(4)
30	3	—	—	3	3
31	2	—	—	2	5
32	2	—	1	1	6
33	—	—	1	−1	5
34	1	2	—	−1	4
35	—	2	—	−2	2
36	—	1	—	−1	1
37	—	—	—	—	1
38	—	1	—	−1	0

This example is merely an illustration, as the figures are obviously too small to be of any use in a real investigation. Nevertheless, it illustrates the points to which attention was drawn in para. 3, viz.

(*a*) E_x finally becomes zero.

(*b*) Equation (2) holds at any stage, e.g.

$$E_{35}=(3+2+2+1)-(2+2)-(1+1)=2.$$

It should be noted that E_x steadily increases with age until a maximum is reached and thereafter decreases. This is a common though not an essential feature.

5. Fractional periods of exposure.

We have shown that an exposed to risk formula can be used when the data are of the simplest type. Let us now see whether the same process can be applied to data of the more complicated type described in the latter part of Chapter II. Two important differences are immediately obvious. The experience has been extended to include the periods between 1st January 1930 and the birthdays of the lives in that year, and between the birthdays in 1934 and 31st December 1934. Moreover, during the investigation period, lives enter and withdraw from the experience. It is still

convenient, however, to deal with the relation between E_x and E_{x-1} under the same headings as in para. 2, viz. beginners, enders and deaths, with the addition of a new heading, new entrants and withdrawals.

(*a*) *Beginners.* The b_x beginners at age x will now be the lives who at the start of the experience on 1st January 1930 were aged between x and $x+1$. As before, let s be the fractional period between the birthday and the end of the calendar year. (There is generally no reason why the fraction s should vary with x since it depends only on the distribution of the birthdays, and we shall assume it to be constant for all ages in the experience.) Leaving out of account for the present the possibility of exit from causes other than death before attaining age $x+1$, we see that all the lives included in b_x were on the average exposed for a fraction $(1-s)$ of a year between ages x and $x+1$, so that we must add $(1-s)\,b_x$ to E_{x-1} to obtain E_x. This adjustment corresponds to (*a*) of para. 2, but a further point now arises. The lives included in b_x—again neglecting exits—were exposed for a full year between ages $x+1$ and $x+2$, and since our expression for E_x allows exposure for only a fraction $1-s$ of a year, we must add sb_x to E_x in calculating E_{x+1}. Arguing on similar lines for b_{x-1}, we see that, so far as terms involving b are concerned, we obtain E_x from E_{x-1} by adding $sb_{x-1}+(1-s)\,b_x$.

(*b*) *Enders.* For enders, our argument is similar. The e_x lives aged between x and $x+1$ on 31st December 1934 were on the average exposed during the investigation for a fractional period s between those ages, neglecting new entrants between ages x and $x+1$. These lives were exposed for a whole year at age $x-1$, again neglecting new entrants. Hence $(1-s)\,e_x$ must be deducted from E_{x-1} to obtain E_x. The lives were not exposed at all between ages $x+1$ and $x+2$ during the investigation period and hence the balance of exposure se_x must be deducted from E_x in finding E_{x+1}. So far, then, as terms involving e are concerned, we obtain E_x from E_{x-1} by deducting $se_{x-1}+(1-s)\,e_x$. The analogy with the terms involving b is obvious.

(*c*) *New entrants and withdrawals.* Let n_x be the number entering between exact ages x and $x+1$ and let $x+k$ be the average

exact age at entry. We shall assume that k is constant at all ages. Although the correctness of this assumption is not so apparent as in para. (a), there is in general no reason why k should vary widely and, as the assumption of a constant value of k for all ages is necessary if a general relation between E_x and E_{x+1} is to exist, we shall take it that variations in k are insufficient to be of practical importance. On the average, then, the new entrants between ages x and $x+1$ will contribute $1-k$ to the exposed to risk between ages x and $x+1$ and, by the same argument as before, the term $kn_{x-1}+(1-k)\,n_x$ should be included in the formula connecting E_x and E_{x-1}.

Similarly, if w_x be the number withdrawing between ages x and $x+1$, the term involving w will be $-hw_{x-1}-(1-h)\,w_x$, where $x+h$ is the average age at the date of withdrawal of lives who withdraw between ages x and $x+1$. (We assume that h, like k, is constant for all ages.)

(d) *Deaths*. In (II, 10) it was pointed out that, in the expression for E_x contained in formula (5), lives dying between ages x and $x+1$ are exposed for a full year, except those beginners and new entrants between ages x and $x+1$ who did not come under review until age $x+t$ $(0<t<1)$, each of whom contributes a fraction $1-t$. We shall take this as our rule in deciding how to deal with deaths in the exposed to risk formula.

Deaths between ages x and $x+1$ during the period of the investigation may be divided into four classes:

(i) Deaths occurring between the birthdays in 1930 and 1934. These should clearly be treated as exposed for a full year between ages x and $x+1$ with no exposure at later ages.

(ii) Deaths occurring in 1930 between ages x and $x+1$ before the birthday in that year.

These will arise out of the b_x lives aged between x and $x+1$ on 1st January 1930 (neglecting for the moment the new entrants who may also contribute to these deaths). On the average, therefore, they should be exposed for a fraction of a year $1-s$ between ages x and $x+1$ and not exposed at all at later ages. The term $(1-s)\,b_x$ added to E_{x-1} provides this exposure for the deaths (as well as for the beginners who do not die), but unless some deduction were

made a similar exposure would be allowed between ages $x+1$ and $x+2$ by the normal carry forward of exposed to risk formulae, and the term sb_x which is added to E_x in calculating E_{x+1} would increase this exposure to unity both for the deaths and the survivors. This is quite correct for the survivors, but the exposure must be cancelled for the deaths by deducting their number from E_x when calculating E_{x+1}.

(iii) Deaths occurring between ages x and $x+1$ in 1934 after the birthday in that year.

These should be exposed for a full year at age x and not exposed at all at older ages, in accordance with the rule given at the beginning of the paragraph.

(iv) Deaths occurring among the new entrants n_x before the attainment of age $x+1$.

These should be exposed, on the average, for a fraction $1-k$ of a year between ages x and $x+1$, no exposure being allowed thereafter. By similar arguments to those advanced in (ii), it can be shown that the required exposure will be allowed by deducting the number of these deaths from E_x in the relation giving E_{x+1}.

Hence it appears that, in calculating E_x from E_{x-1}, no deduction should be made for any of the deaths between ages x and $x+1$, but that a unit should be deducted for each of them in the relation connecting E_{x+1} and E_x, so that θ_{x-1} should be deducted from E_{x-1} in calculating E_x and θ_x from E_x in calculating E_{x+1}.

Finally, therefore, combining the results of sections (*a*), (*b*), (*c*) and (*d*) of this paragraph, we obtain the exposed to risk formula:

$$\begin{aligned} E_x = E_{x-1} + sb_{x-1} + (1-s)\, b_x - se_{x-1} - (1-s)\, e_x \\ + kn_{x-1} + (1-k)\, n_x - hw_{x-1} - (1-h)\, w_x - \theta_{x-1}. \quad \ldots(2) \end{aligned}$$

6. Test of formula.

It may now be instructive to trace the passage of individual lives through the investigation. Lives may enter the experience either as beginners or new entrants and may leave as enders, withdrawals or deaths. Each of the two methods of entering can be combined with any of the methods of leaving, making six combinations in

all, which we shall distinguish by the appropriate pair of symbols. Let us consider each case separately and verify that our formula deals with it correctly.

(*a*) *b and e.* If the life is aged between x and $x+1$ on 1st January 1930, we add $1-s$ to E_{x-1} in calculating E_x, and an additional fraction s, making a unit in all, in calculating E_{x+1}. This unit is carried foward into the exposed to risk at older ages. On 31st December 1934 the life will be between ages $x+5$ and $x+6$ and will be included in e_{x+5}. A fraction $1-s$ will therefore be deducted from E_{x+4} in calculating E_{x+5}, leaving a balance of exposure s between ages $x+5$ and $x+6$, and this fraction s will be deducted from E_{x+5} in calculating E_{x+6}, thus removing the life from the exposed to risk at $x+6$ and older ages.

(*b*) *b and w.* The effect of including a unit in b_x will be the same as in (*a*). If withdrawal occurs before age $x+1$, a unit will be included in w_x and as a result a deduction of $1-h$ will be made in calculating E_x from E_{x-1}, with a further deduction of h in calculating E_{x+1}. On balance, an exposure of $h-s$ will be included in E_x with no exposure at later ages.

If withdrawal occurs between ages $x+1$ and $x+2$, there will be no deduction in calculating E_x, but $1-h$ will be deducted in calculating E_{x+1} and an additional h in calculating E_{x+2}. On balance, the exposure allowed will be $1-s$ at age x, h at age $x+1$ and nothing thereafter; and similarly for withdrawals at later ages.

(*c*) *b and θ.* The effect of including a unit in b_x is again the same as in (*a*). If death occurs between ages x and $x+1$ a unit will be included in θ_x and a deduction of unity will be made in calculating E_{x+1} from E_x, so that the exposure allowed will be $1-s$ at age x and nothing thereafter. If death occurs at age $x+1$ the exposure will be $1-s$ at age x, unity at age $x+1$ and nothing thereafter. Finally, if death occurs at age $x+5$, i.e. between the attainment of that age in 1934 and the end of the year, the exposure will be $1-s$ at age x, unity at ages $x+1$ to $x+5$ inclusive and nothing thereafter.

(*d*) *n and e.* For a new entrant at age x, an exposure of $1-k$ will be included in E_x and an additional k in E_{x+1}, so that the exposure will be $1-k$ at age x and unity at age $x+1$, and this unit

will be carried forward to later ages. If the life is included as an exit in, say, e_{x+3}, the exposures can be worked out by the same process as in (*a*). In the special case where entry took place after the birthday in 1934, the exposure included at age x, the age attained in 1934, will be $s-k$, with no exposure at older ages.

(*e*) *n and w.* This can be investigated in the same way as (*b*). The exposure will be $h-k$ at age x, if entry and withdrawal occur in the same year of age; otherwise, it will be $1-k$ in the year of age at which the life entered, h in the year of age at which it withdrew and unity at intermediate ages.

(*f*) *n and θ.* The same argument can be applied as that used in (*c*).

It will be seen that in every case the exposure allowed by the exposed to risk formula is the same as that allowed by formula (5) of Chapter II. The reader should check this for each of the six cases.

We have now established a method of linking the exposed to risk at age x with that at age $x-1$ which will give exactly the same results as those obtained by the more laborious method of the last chapter. It will be noticed that in constructing the exposed to risk formula we considered the terms relating to entrants without regard to the incidence of the exits and vice versa. If we consider the conception of the exposed to risk as the difference between the numbers who entered and the numbers who passed out of observation at ages up to x (para. 3), the analogy is apparent, but we are now dealing with years of life instead of with lives, since fractional exposures have been introduced.

7. Alternative demonstration of the treatment of fractional exposures.

It is interesting and instructive to approach the question of fractional exposures from a different angle. Let us consider for example the b_x beginners who at the start of the experience are aged $x+s$ on the average. They contribute $(1-s)\,b_x$ to E_x and $(1-s)\,b_x+sb_x(=b_x)$ to E_{x+1}, E_{x+2}, etc. Clearly the same effect would be produced by a group of $(1-s)\,b_x$ lives aged x exactly and

another group of sb_x lives aged $x+1$ exactly at the start of the experience. Similarly, the beginners at age $x-1$ give us $(1-s)\,b_{x-1}$ at exact age $x-1$ and sb_{x-1} at exact age x. We may therefore regard the beginners as consisting of groups of $(1-s)\,b_x+sb_{x-1}$ lives entering the experience at exact age x. By repeating this process, we obtain similar expressions for enders, new entrants and withdrawals and in this way we reduce the problem to the simple case considered in para. 2, where all movements except deaths take place at exact ages. Formula (1) of para. 2, when extended to allow for new entrants and withdrawals, expresses the relationship

$$\begin{aligned} E_x = E_{x-1} &+ (\text{beginners at exact age } x + \text{new entrants at exact age } x) \\ &- (\text{enders at exact age } x + \text{withdrawals at exact age } x \\ &\qquad + \text{deaths at age } x-1 \text{ last birthday}). \end{aligned}$$

If we substitute in this formula the expressions obtained above we arrive at once at formula (2) of para. 5.

The reader will understand that the demonstration in this paragraph does not involve a method essentially different from that used previously. The assumptions involved are exactly the same as before and the results are identical. The methods differ only in the route followed in proceeding from one to the other.

8. The demonstration of exposed to risk formulae given in this chapter has necessarily involved a good deal of detail and it is hoped that this has not obscured the underlying principles. Nothing in the chapter need or should be memorized. If the reader has understood clearly the basis of the method explained in paras. 2 and 3 he will have taken a definite forward step in his study of the subject of construction of tables and he should have no difficulty in drawing up a formula on the lines of formula (2) by translating principles into practice. Without such an understanding, no amount of memory work will be of any avail.

One fruitful source of error is a confusion of the two variables time and age. An exposed to risk formula shows the relation between the exposed at successive ages. The element of time is not directly involved, but for each calendar year passed through

there is an increase in age and we are therefore tempted to relate movements to the calendar year in which they occur and not to the year of life. In this way confusion arises. It is therefore advisable to relate all movements to the year of *life* rather than to the year of *time* or calendar year in which they take place. It is important, too, when using the word "years" to make it clear whether the reference is to years of life or years of time.

CHAPTER IV

EXPOSED TO RISK FORMULAE. AGE GROUPING BY DATE

1. In the last two chapters we discussed the problem of calculating rates of mortality under what may be called ideal conditions, for we assumed that we can obtain for each case included in the experience complete and exact information which will enable us to determine accurately the ages at which the life enters and leaves the investigation. We traced the lives from birthday to birthday and for this reason the method is sometimes called the *life-year* method.

In practice, full information may not be available in a convenient form. It is true that if we know the dates of birth we can calculate the ages required, but in a large experience this will involve a considerable amount of work which may be avoided by using the data in the form in which they are most readily available. Moreover, it sometimes happens that the dates of birth are not known. In such cases the methods of Chapter III are unsuitable and we must seek other ways of arriving at exposed to risk formulae. These invariably involve some method of age grouping other than by exact age, e.g. by age last birthday, nearest age or calendar year of birth. In other words, the lives are grouped by approximate age instead of by exact age.

2. Approximations to age.

Broadly speaking there are two ways of approximating to age:

(i) We can take our stand at a fixed date in the calendar year and group together all lives whose ages at that date lie within certain limits. This group must then be traced until the fixed date in the following year. We may, for example, take our fixed date as 1st January and group together all lives who are aged x nearest birthday on any 1st January within the period of the investigation. Having fixed our age group in this way, we then proceed to trace the lives included therein throughout the succeeding calendar year.

The calendar year in question is not the same for all the lives unless the experience is limited to a single calendar year, for we have included in the same group all lives aged x nearest birthday on any 1st January included in the experience. (There is an obvious analogy with the method of Chapter III in which we grouped together all lives attaining exact age x at any time within the investigation period.) Lives aged x nearest birthday on one 1st January will, if they survive, be aged $x+1$ nearest birthday on the following 1st January and will contribute to the experience for that age also, provided that the latter date is still within the limiting dates of the experience.

This method of approximating to age is suitable when the data are most readily available in the form of the valuation records of a life office or a friendly society, for in these records the lives will usually be grouped according to their approximate ages on 31st December or some other fixed date.

(ii) Alternatively, we may base our grouping not on the ages at a fixed date but on the ages on the anniversary of a certain event. In practice, the event chosen is generally one that brings the life into the experience or alters its exposed to risk status in some way. Typical "events" are retirement in a pension fund experience, marriage or widowhood in a widows' fund, disability in a sickness or group life investigation and, most common and obvious of all, the effecting of a life assurance or annuity contract in life office work.

In these cases our records will almost certainly yield information about the age in a convenient form which will enable us to group together all lives who were, let us say, aged x next birthday on the anniversary of the date on which the policy was effected. Lives in the group so formed must then be traced up to their next policy anniversaries.

For convenience, we shall refer to these two methods of approximation as grouping by date and grouping by event respectively and, since the derivation of exposed to risk formulae depends to some extent on the grouping we select, we shall deal with the two separately. In this chapter various methods of grouping by date will be discussed.

3. Method I—Age nearest birthday.

Let us begin with the case where the office valuation or other records give the nearest age on the valuation date in any year. For simplicity, let us take this date to be 1st January and assume that our investigation covers only one calendar year.

In Chapter III we considered a group of lives aged x exactly and traced them through the succeeding year of age to exact age $x+1$, taking account of new entrants, deaths and withdrawals between those ages and hence calculating the exposed to risk and rate of mortality at exact age x. There is no reason why a similar process should not be followed for a group of lives who, instead of being aged x exactly, are aged x nearest birthday.

If there are no new entrants or withdrawals during the year, the exposed to risk will be the number of lives aged x nearest birthday on 1st January and the corresponding deaths will be those occurring during the calendar year *among these lives*. Clearly we must not take the deaths occurring between *exact* ages x and $x+1$ during the calendar year as we did in the previous method, since some of those dying between exact ages x and $x+1$ would not be aged x nearest birthday on 1st January and, on the other hand, some of the deaths during the calendar year among the lives under review would not occur between exact ages x and $x+1$. We would therefore violate the fundamental principle that all the deaths occurring during the year among the lives exposed to risk must be included *and no other deaths*.

This method gives the rate of mortality among lives aged x nearest birthday on a particular 1st January. It may be extended to an investigation covering several calendar years and the exposed to risk at age x will then be the number of lives whose nearest age on any 1st January during the period of the investigation was x. To make θ_x and E_x consistent, the deaths during the period of the investigation must be classified according to age nearest birthday on the 1st January preceding death.

New entrants and withdrawals, which have so far been neglected, may be dealt with similarly and will also be classified under age nearest birthday on the 1st January preceding entry or withdrawal.

It should be noted that the method does not exclude any of the

available data, for every life on the books of the office in a particular calendar year will fall into one or other of the groups having the general age description "x nearest birthday".

4. Method I—Exposed to risk formula.

We can now proceed directly to construct the exposed to risk formula for the general case of an investigation covering a period of several years. As before, we will take the period from 1st January 1930 to 31st December 1934. The symbols to be used will be defined as follows:

b_x = number of beginners aged x nearest birthday on 1st January 1930.

e_x = number of enders aged x nearest birthday on 31st December 1934.

n_x = number of new entrants in the years 1930–34 aged x nearest birthday on 1st January prior to entry.

w_x = number of withdrawals in the years 1930–34 aged x nearest birthday on 1st January prior to withdrawal.

θ_x = number of deaths in the years 1930–34 aged x nearest birthday on 1st January prior to death.

For each of the b_x beginners a unit must be included in E_x but nothing in E_{x-1}, and for each of the e_x enders a unit in E_{x-1} but nothing in E_x. For each of the θ_x deaths we must deduct a unit in calculating E_{x+1} from E_x as in (III, 5 (d)). For new entrants and withdrawals, let k and h be the average periods between the dates of entry and withdrawal respectively and the preceding 1st January. Then, by the argument of (III, 5), the exposed to risk formula is

$$E_x = E_{x-1} + b_x - e_x + kn_{x-1} + (1-k)\,n_x - hw_{x-1} - (1-h)\,w_x - \theta_{x-1}. \qquad \text{......(1)}$$

It should be noted that, although E_x is referred to as "the exposed to risk at age x", this does not necessarily mean exact age x but the age x determined by the definitions of the symbols. Here, for example, x means the age nearest birthday on 1st January; the rate of mortality is therefore that applicable to a body of lives aged x nearest birthday on any 1st January in the years 1930–34. To see what this means in terms of exact ages, we must consider in more detail the ages of the lives included in the exposed to risk.

5. Method I—Analysis of exposed to risk.

Lives aged x nearest birthday on 1st January are between exact ages $x-\frac{1}{2}$ and $x+\frac{1}{2}$ on that date and between exact ages $x+\frac{1}{2}$ and $x+\frac{3}{2}$ on the following 31st December. In the exposed to risk of the previous paragraph we therefore have lives of all exact ages between $x-\frac{1}{2}$ and $x+\frac{3}{2}$. We shall assume that birthdays are uniformly distributed over the calendar year and, for the sake of simplicity, that there are no entrants and withdrawals during the year. Then if there are N lives aged x nearest birthday on 1st January the number between exact ages $x+t$ and $x+t+\Delta t$, where Δt is small, will be $N\Delta t$ for all values of t $(-\frac{1}{2}\leqslant t\leqslant\frac{1}{2})$.

(*a*) If $-\frac{1}{2}\leqslant t\leqslant 0$, each of the group of $N\Delta t$ lives will contribute $-t$ to the exposed to risk between ages $x-\frac{1}{2}$ and x and $(1+t)$ to the exposed to risk between ages x and $x+1$. The total exposed to risk is thus

$$N\int_{-\frac{1}{2}}^{0}(-t)\,dt=N/8, \text{ between ages } x-\tfrac{1}{2} \text{ and } x,$$

and

$$N\int_{-\frac{1}{2}}^{0}(1+t)\,dt=3N/8, \text{ between ages } x \text{ and } x+1.$$

(*b*) If $0\leqslant t\leqslant\frac{1}{2}$, each of the group of $N\Delta t$ lives will contribute $(1-t)$ to the exposed to risk between ages x and $x+1$ and t to the exposed to risk between ages $x+1$ and $x+\frac{3}{2}$. The total exposed to risk is thus

$$N\int_{0}^{\frac{1}{2}}(1-t)\,dt=3N/8, \text{ between ages } x \text{ and } x+1,$$

and

$$N\int_{0}^{\frac{1}{2}}t\,dt=N/8, \text{ between ages } x+1 \text{ and } x+\tfrac{3}{2}.$$

Hence the exposure for one year contributed by N lives aged x nearest birthday at the beginning of the year of exposure is made up as follows:

$N/8$ years of exposure between ages $x-\frac{1}{2}$ and x,
$3N/4$ " " " x and $x+1$,
$N/8$ " " " $x+1$ and $x+\frac{3}{2}$.

Thus the greater part of the exposed to risk is strictly applicable

to the year of age x to $x+1$ and, in general, the rate of mortality obtained for lives aged x nearest birthday on 1st January will be very nearly the rate of mortality for exact age x.

Even if birthdays are not uniformly distributed throughout the year, the method does not fail. We still arrive at a rate of mortality for nearest age x, only now the average exact age of the lives will not be x but $x+l$, where $-\frac{1}{2}<l<\frac{1}{2}$ and the rate of mortality will be q_{x+l} and not q_x. The value of l can be obtained by examining the dates of birth of the lives, and hence values for $q_x, q_{x+1} \cdots$ can be deduced by interpolation from the table of $q_{x+l}, q_{x+1+l} \cdots$.

The fact that nearest age x is not necessarily an approximation to exact age x emphasizes the importance of remembering that the symbols E_x, b_x, etc. have no fixed meaning and must be defined before they are used in any demonstration. This can be seen by comparing the meanings assigned to them in para. 4 with their meanings in formulae (1) and (2) of Chapter III.

6. Method II—Age nearest birthday on a fixed date other than the starting date of the experience.

Formula (1) was simplified by the fact that the starting date of the investigation coincided with the date in the year used for purposes of age grouping. Cases may occur, however, where it will be convenient to group the lives according to nearest age on some date in the year other than the starting date of the experience or its anniversary. Suppose that, as before, our investigation covers the period from 1st January 1930 to 31st December 1934 and that the lives are grouped according to nearest age on 1st April. E_x will then consist of the number of years of exposure in the years 1930–34 contributed by lives in the year of life following the 1st April on which they were aged x nearest birthday. The definitions of n_x, w_x and θ_x and of k and h in para. 4 will be amended by substituting 1st April for 1st January and the other symbols will be defined as follows:

b_x = number of beginners on 1st January 1930 who were aged x nearest birthday on 1st April 1929,

e_x = number of enders on 31st December 1934 who were aged x nearest birthday on 1st April 1934.

The b_x beginners come under observation at assumed age $x+\frac{3}{4}$ and $\frac{1}{4}$ must therefore be included in E_x for each of them. They are exposed for a full year at assumed age $x+1$, so that a further $\frac{3}{4}$ must be added in obtaining E_{x+1} from E_x. A similar argument holds for the e_x enders and the exposed to risk formula accordingly becomes:

$$E_x = E_{x-1} + \tfrac{3}{4}b_{x-1} + \tfrac{1}{4}b_x - \tfrac{3}{4}e_{x-1} - \tfrac{1}{4}e_x \\ + kn_{x-1} + (1-k)\,n_x - hw_{x-1} - (1-h)\,w_x - \theta_{x-1}. \quad \ldots(2)$$

The phrase "assumed age $x+\frac{3}{4}$" is used because it would be meaningless to apply the words "nearest age" to anything except an integral age. Moreover, it provides a convenient description which can be applied in the case of any method of age grouping to avoid repeating a full explanation of the method employed every time an age is mentioned. It does not necessarily imply an approximation to exact age x; it may, for example, be used when lives are grouped according to age next birthday.

7. Method III—Age next (or last) birthday.

If the grouping is by age next birthday instead of age nearest birthday on 1st January, we can repeat the argument of paras. 3 and 4, substituting age next birthday for age nearest birthday throughout. The resulting rate of mortality will apply to age x next birthday, i.e. it will be the rate for an exact age less than x. If we analyse the exposed to risk as was done in para. 5, we shall find that the exposure contributed by a group of N lives aged x next birthday at the beginning of the year of exposure is made up as follows:

$N/8$ years of exposure between ages $x-1$ and $x-\frac{1}{2}$,
$3N/4$ " " " $x-\frac{1}{2}$ and $x+\frac{1}{2}$,
$N/8$ " " " $x+\frac{1}{2}$ and $x+1$,

and that therefore the rate of mortality will be a close approximation to the rate for exact age $x-\frac{1}{2}$.

Similarly, if our grouping is by age last birthday, we shall obtain the rate for age $x+\frac{1}{2}$.

In both cases the exposed to risk formula will be the same as formula (1), but the meaning of the symbols will of course be different since we have substituted age x next (or last) birthday for age x nearest birthday in the definitions.

If the birthdays are not uniformly spread over the calendar year, the equivalent exact age will be $x-l$ for age next birthday and $x+l$ for age last birthday, where $0 \leqslant l \leqslant 1$ and the value of l depends on the distribution of the birthdays.

8. Method III—Practical application.

This method of age grouping is of special importance in practice. It possesses the great advantage that it can be carried out without regard to the dates of birth for, if we deduct the calendar year of birth from the calendar year in which an event occurs, we obtain in all cases the age next birthday on 1st January preceding the event. Thus, a life born in 1910 who leaves the experience in 1933 is aged 23 next birthday on 1st January 1933. (If we are using age last birthday we deduct one *plus* the calendar year of birth from the year of the event.)

This is a considerable saving of work. To find, for example, the age nearest birthday on 1st January in any year, we must first look to see whether the birthday occurs before or after 30th June so as to obtain the nearest birthday and then calculate the age at the birthday in question. It is obviously a simpler process to obtain the value of assumed age x merely by subtracting year of birth from year of entry or exit.

If, as sometimes happens, only the calendar year of birth is recorded and the exact date of birth is unknown, the present method is the only one which can be easily applied.

Let us take as before an experience covering the years 1930–34 inclusive and assume that the calendar years of birth, entry and exit are recorded for each life coming into the investigation. The meaning to be assigned to x will then be:

for beginners: 1930 − calendar year of birth (i.e. b_x consists of beginners on 1st January 1930 who were born in the calendar year $1930-x$);

for enders: 1935 − calendar year of birth;

$$\text{for} \left\{\begin{array}{l}\text{new entrants:}\\ \text{withdrawals:}\\ \text{deaths:}\end{array}\right\} \text{calendar year of} \left\{\begin{array}{l}\text{entry}\\ \text{withdrawal}\\ \text{death}\end{array}\right\} - \begin{array}{l}\text{calendar}\\ \text{year of}\\ \text{birth.}\end{array}$$

As already explained, these age descriptions are in effect exactly the same as those in para. 4, if age next birthday is substituted for age nearest birthday in the definitions of that paragraph.

9. The groupings dealt with in the preceding paragraphs illustrate various methods of approach. The principles on which they are based can with suitable modifications of detail be applied to any other cases which arise. It is therefore unnecessary to multiply examples. It may, however, be of assistance to add two which are typical of conditions occurring in practice where the information available is rather less complete than we have hitherto assumed.

10. Method IV—Age nearest birthday at entry and calendar years of entry and exit.

We have seen that method III can be applied successfully if we know the calendar years of birth, entry and exit. Suppose that instead of the calendar year of birth we know the age nearest birthday at entry. We cannot then determine accurately the nearest age or the age next or last birthday on the preceding 1st January (or on any other fixed date) as we have been able to do in previous methods.

For example, a life entering in 1932 at age 25 nearest birthday may have entered on 1st January 1932—in which case his date of birth lies between 1st July 1906 and 30th June 1907—or on 31st December 1932—in which case his date of birth lies between 1st July 1907 and 30th June 1908. The age on 1st January 1932 of such an entrant may therefore lie anywhere between exact ages $23\frac{1}{2}$ and $25\frac{1}{2}$. If we assume that birthdays and dates of entry are uniformly spread over the calendar year, the average age of lives in this group on 1st January 1932 will be $24\frac{1}{2}$ and the average date of birth 30th June 1907, i.e. we may take the calendar year of birth to be 1907 on the average or, in the general case, calendar year of entry *minus* nearest age at entry.

This method of grouping is similar to method III above and we may adopt the age groupings given in paragraph 8, substituting for the true calendar year of birth the estimated calendar year of

birth, i.e. calendar year of entry − age nearest birthday at entry; this is usually known as *office year of birth*. We define x thus:

for beginners: 1930 − (office year of birth);
for enders: 1935 − (office year of birth);
for new entrants: age nearest birthday at entry;
for withdrawals: calendar year of withdrawal } − (office year of
deaths: death } birth).

In practice, we would facilitate calculation of the assumed ages of the movements by recording the office year of birth for each life. This will not necessarily be the true year of birth, but will be treated as such in obtaining the ages of the various movements. Once these ages have been obtained, the practical details of the method are identical with those of method III above and the resulting rate of mortality is applicable to exact age $x-l$, where $0 \leqslant l \leqslant 1$.

In the absence of any information about the distribution of entry dates and dates of birth over the calendar year, we should probably assume an even distribution in each case which would give to l the value of $\frac{1}{2}$.

11. Method V—Age next birthday at entry and calendar years of entry and exit.

If we assume that, as before, birthdays and dates of entry are uniformly spread over the calendar year, then a group of lives aged x next birthday at entry will be on the average half a year younger at entry than lives aged x nearest birthday at entry. This suggests that by proceeding exactly as in method IV we shall obtain a rate of mortality not for age $x-\frac{1}{2}$ but for age $x-1$. To confirm that this is so, we need only remark that for a group of lives entering in a particular calendar year the next birthday after entry will on the average fall on 1st January in the next calendar year. The average nearest age on 1st January before entry of a group of lives aged x next birthday at entry is thus $x-1$. In method IV we saw that this was $x-\frac{1}{2}$.

The exposed to risk resulting from the age descriptions of para. 10 with age next birthday substituted for age nearest birthday throughout leads therefore to an approximation to the rate of mortality for age $x-1$. Similarly, if we used age last birthday at

entry instead of age next birthday, we should arrive at a rate of mortality for age x. In either case, the correctness of the assumptions should if possible be tested.

12. Methods IV and V—Analysis of exposed to risk.

It is evident that in methods IV and V we have had to be content with a wider age grouping and a less accurate approximation to the age. Consider, for example, two lives who contribute to E_{24} when we are grouping according to age last birthday at entry plus calendar year of movement less calendar year of entry.

(*a*) A life born on 2nd January 1908 who entered on 1st January 1930, age last birthday at entry 21, will contribute to E_{24} in 1933 for, taking the age description of method V, we have

$$x = 1933 - (1930 - 21) = 24.$$

The exact age on 1st January 1933 is however 25.

(*b*) A life born on 30th December 1906 who entered on 31st December 1928, age last birthday at entry 22, will contribute to E_{24} in 1930, for $\quad x = 1930 - (1928 - 22) = 24.$

The exact age on 1st January 1930 is however 23.

These are two extreme cases showing the widest possible variation in the ages at the beginning of the year of exposure of lives assumed to be age 24. The corresponding ages at the end of the year of exposure will lie between 24 and 26, so that the whole range of exposure is from 23 to 26 instead of from 24 to 25.

It is interesting to analyse the exposures for this case as was done in para. 5. The analysis is more involved and need not be set out in detail, but we may note here the result, namely that the exposed to risk is distributed thus:

$\frac{1}{6}$th between ages $x-1$ and x,
$\frac{2}{3}$rds ,, ,, x and $x+1$,
$\frac{1}{6}$th ,, ,, $x+1$ and $x+2$.

The range of ages covered is thus three years instead of two as in method I and the proportion of the exposed to risk falling within the year x to $x+1$ is $\frac{2}{3}$ compared with $\frac{3}{4}$ in method I. Clearly therefore the resulting rate will not in general be so good an approximation as by the previous method.

When nearest age at entry or age next birthday at entry is given, the distribution of the exposed to risk will be similar but will extend from age $x-1\frac{1}{2}$ to age $x+1\frac{1}{2}$ and from age $x-2$ to age $x+1$ respectively.

13. Various references have been made above to the distribution of the exposed to risk at assumed age x over a range of several years of age. It is unnecessary and may be confusing to keep this idea too prominently in mind during the actual operations, for once the age grouping has been settled and the assumed ages of entering and leaving recorded for each life, the exposed to risk formula can be drawn up just as if assumed age x were the same as exact age x. It is of course essential that, once this has been done, we should investigate the age distribution of the exposed to risk to decide the approximate exact age to which the rate of mortality for assumed age x corresponds.

14. Methods of deriving exposed to risk formulae of the type described in this chapter are usually called *calendar year* methods. The phrase emphasizes the fact that we trace the progress of all the lives from a fixed date in one year to the same date in the next and subsequent years—frequently from 1st January to 1st January, i.e. by calendar year. It ignores however the important element of age. What we have really done is to choose one day of the year on which to approximate to the ages of the lives and we trace the lives from year to year only because we thereby trace them from one assumed age to the next.

To keep the assumed age rather than the calendar year prominently in front of us is of great practical importance when we are considering movements, for our aim is to assign them to the correct assumed ages at which they occur. The fact that a movement took place in a certain calendar year is in itself of small importance, but it is important to know that on the preceding 1st January the life was of assumed age x and therefore that the movement belongs to the year of age x to $x+1$. For this reason, it seems better to describe the methods of this chapter as examples of the method of age grouping by date.

CHAPTER V

EXPOSED TO RISK FORMULAE. AGE GROUPING BY EVENT

1. We now proceed to consider the second type of age approximations referred to in (IV, 2), namely those based on *event*. The most important example is that of a life office, where the terms of life assurance and annuity contracts are governed by the age at the date of entry, so that this age is recorded and is conveniently available. The age on any policy anniversary can be obtained by adding to the age at entry the difference between the calendar year in which the anniversary falls and the calendar year of entry.

This suggests that E_x might be based on the lives attaining assumed age x on any policy anniversary during the period of the investigation. To secure consistency between the deaths and the exposed to risk, θ_x must then be the number of lives dying during the period of the investigation who attained assumed age x on the policy anniversary before death.

All movements must similarly be grouped according to the assumed age at the beginning of the *policy* year in which the movement occurred, just as by the method of Chapter II they were grouped according to the age at the beginning of the *life* year in which the movement occurred. This may be done by recording for each case the duration since entry on the policy anniversary preceding the date of coming under observation and the corresponding duration on passing out of observation. These are usually described as the *curtate durations* at which the movements occur and they are, of course, the next lower integral durations to the exact durations. For example, if a life entered on 1st April 1925 and came under observation on 1st January 1930, the exact duration on the latter date was $4\frac{3}{4}$ years and the curtate duration 4 years.

The analogy to the methods of Chapters II and III is easily seen, for, instead of including the lives attaining exact age x on any birthday during the period of the investigation, we include the

lives who were of assumed age x on any policy anniversary during that period. The exposed to risk formulae can therefore be constructed in the same way as in Chapter III. It may, however, be instructive to describe the process in detail.

2. Method I—Policy anniversaries as limiting dates of experience.

Let us consider as before an investigation covering the years 1930–34, but extending only from the policy anniversaries in 1930 to the policy anniversaries in 1934. This is analogous to the method of (III, 2), where the limiting dates of the investigation were the birthdays in 1930 and 1934. We have then the following definitions:

b_x is the number of beginners on their policy anniversaries in 1930, such that x equals the assumed age at entry plus the exact duration on the policy anniversary in 1930.

e_x is the number of enders on their policy anniversaries in 1934, such that x equals the assumed age at entry plus the exact duration on the policy anniversary in 1934.

n_x is the number of new entrants at assumed age x.

θ_x and w_x are the numbers of deaths and withdrawals such that x equals the assumed age at entry plus the curtate duration at exit.

New entrants in 1934 and deaths and withdrawals which take place before their policy anniversaries in 1930 will be excluded from their investigation. Deaths and withdrawals after their policy anniversaries in 1934 will be treated as enders.

If h is the average period between the date of withdrawal and the preceding policy anniversary, the exposed to risk formula is

$$E_x = E_{x-1} + b_x - e_x + n_x - hw_{x-1} - (1-h)\, w_x - \theta_{x-1}. \quad \ldots\ldots(1)$$

b_x and n_x both come under review at assumed age x and therefore contribute a whole unit each at that age, while e_x contribute a whole unit each at assumed age $x-1$ but nothing at assumed age x. w_x are each exposed for a fraction h at assumed age x and not thereafter.

In practice, assumed age x will generally be either age next birthday for assured lives or age last birthday for annuitants, since these are the ages for which life assurance premiums and annuity

rates are most commonly quoted. There is a tendency for lives to effect life assurance policies shortly before a birthday and there is also a tendency to purchase annuities shortly after a birthday. The exact ages corresponding to age x next birthday and age x last birthday will thus lie between $x-\frac{1}{2}$ and x and between x and $x+\frac{1}{2}$ respectively. To determine the precise values it would be necessary to investigate the age distribution of the data, but in the absence of the information required for such an investigation we should probably use the value obtained from a previous experience of the same type. The equivalent exact age has usually been found to be in the region of $x-\frac{1}{3}$ for assurances and $x+\frac{1}{3}$ for annuities.

3. Method II—Fixed limiting dates.

Still following the lines of Chapter III, we may extend our investigation to cover the whole period from 1st January 1930 to 31st December 1934. The definitions of para. 2 will require to be amended thus:

b_x is the number of beginners on 1st January 1930 such that x equals the age at entry plus the curtate duration at 1st January 1930.

e_x is the number of enders on 31st December 1934 such that x equals the age at entry plus the curtate duration at 31st December 1934.

n_x is the number of new entrants at age x in the years 1930–34.

θ_x and w_x are the numbers of deaths and withdrawals in the years 1930–34 such that x equals the age at entry plus the curtate duration at exit.

s is the average period between 1st January 1930 and the preceding policy anniversary (which we assume to be the same as the average period between 31st December 1934 and the preceding policy anniversary).

h is the average period between the date of withdrawal and the preceding policy anniversary.

The exposed to risk formula can then be constructed in the usual way. The b_x beginners are on the average exposed for a fraction $(1-s)$ of a year between assumed ages x and $x+1$ and

exposed for a whole year between $x+1$ and $x+2$. The appropriate term in the formula for E_x is accordingly $sb_{x-1}+(1-s)\,b_x$. Similar arguments apply for the other movements. The exposed to risk formula is therefore

$$E_x = E_{x-1} + sb_{x-1} + (1-s)\,b_x - se_{x-1} - (1-s)\,e_x + n_x \\ - hw_{x-1} - (1-h)\,w_x - \theta_{x-1}. \quad \ldots(2)$$

As before, the data will have to be investigated to determine the exact age corresponding to the assumed age.

4. From what has been said above and from the close similarity of the formulae to those obtained in Chapter III, it is evident that the grouping of ages by event introduces little that is new in principle. As already mentioned, life office records lend themselves to the application of this method and, for this reason, it is often called the *policy year* method. As we shall see, the method is of considerable importance when we come to investigate mortality as a function not of the age only but also of the duration, which in this context denotes the period which has elapsed since the happening of some event. The method is clearly well adapted for such investigations.

5. The reader may have some doubts about the inclusion of methods IV and V of Chapter IV as examples of age approximations by date, for the use of the age at entry suggests at first sight that they would more correctly be treated as examples of the method of age grouping by event. The truth is that the age at entry is used merely as a part of the process of approximating to the age at the beginning of the calendar year of exposure. The only data necessary for methods IV and V are the age at entry and the calendar years of the movements, whereas the policy year method requires the curtate duration of the movements or else the exact dates in order that they may be grouped according to curtate duration. To sum up, the methods of Chapter IV involve the grouping of lives having a common age description at the beginning of the calendar year of exposure, while those of Chapter V bring together lives having a common age description at the beginning of the policy year of exposure.

6. The use of the terms life, calendar and policy year to classify the various methods is convenient and is now firmly established. The nomenclature is, however, unsatisfactory in that it does not emphasize the essential differences which consist, as we have seen, in the methods of bringing together lives having a common age description. For this reason, the nomenclature adopted in this and the preceding chapter is to be preferred.

More generally, exception may be taken to the use of any standardized nomenclature, because it may suggest that the choice of a method for a particular set of data can be made by a fixed rule and that the procedure thereafter will follow automatically. In actual practice, one method may be best suited to some of the data and another more suitable for the remainder. None of the methods hitherto described can then be applied without further approximations, and careful thought is necessary to decide which system of age grouping will reduce errors to a minimum, i.e. will give the closest relation between deaths and exposures with the available data.

Although no fixed rule can be laid down to determine the choice of method, there is one guiding rule which is generally applicable. This is to fix the method of grouping according to the way in which the deaths are grouped; e.g. if the deaths are grouped according to age next birthday on the previous policy anniversary, a policy year method will be used and assumed age x taken as age x next birthday. The reason is clear when we remember that the numerator of the fraction giving q_x is based entirely on terms in θ, so that an error in approximating to the number of deaths is of greater moment than a similar error in any or all of the terms b, n, etc., each of which forms only a part of the denominator.

This is true even if the chosen method involves approximations to the exposure at assumed age x for all lives subject to movements other than death between assumed ages x and $x+1$. At first glance this may appear contradictory, but it must be remembered that a considerable part of E_x relates to lives who were not involved in any movement at assumed age x and this part will usually be free from errors of the type in question. The reader will better appreciate the argument after studying example 4 of Chapter VII.

The effect of the rule is that the numerator of q_x will contain a single term in θ and not the sum of two terms such as $\frac{1}{2}(\theta_{x-1}+\theta_x)$. If data are available to permit of regrouping to achieve complete consistency, this may be the best solution. The rule is intended for use when regrouping is impossible or inconvenient.

7. Alternative method for fractional exposures.

We shall conclude the chapter by describing an alternative method of dealing with fractional exposures which can be applied in all cases where we know the exact dates on which movements occur. In (III, 7), we saw that, in an exact age method, movements which take place on the average at exact age $x+s$ can be divided into two groups which are allocated to exact ages x and $x+1$ respectively. Thus b_x beginners at age $x+s$ are equivalent to $(1-s)\,b_x$ beginners at age x and sb_x at age $x+1$.

Suppose now that we consider each beginner individually and classify him according to his age nearest birthday; i.e. if his exact age is $x+s$, he is called a beginner at age x when $0<s\leqslant\frac{1}{2}$ and a beginner at age $x+1$ when $\frac{1}{2}<s\leqslant 1$. In this way all the movements can be treated as occurring at integral ages and fractional exposures involving the fractions s, k or h will be avoided. The deaths must still be grouped according to their ages at the birthday preceding death. This is necessary because the numerator of q_x must consist of the deaths between ages x and $x+1$ and the denominator must treat all these deaths as exposed up to age $x+1$.

By an obvious extension, the method can be applied when ages are grouped by date or by event. We shall then use assumed ages instead of exact ages and classify the movements according to the assumed age on the nearest 1st January or the nearest policy anniversary. The deaths will be recorded according to their assumed ages at the beginning of the calendar or policy year as the case may be. The effect of this procedure is to treat movements at assumed age $x+t$ as if they occurred at assumed age x when $0<t\leqslant\frac{1}{2}$ and at assumed age $x+1$ when $\frac{1}{2}<t\leqslant 1$.

The formula for E_x takes the following simple form whether the symbols relate to exact or to assumed ages:

$$E_x=E_{x-1}+b_x-e_x+n_x-w_x-\theta_{x-1}. \qquad \ldots\ldots(3)$$

The resulting values of E_x are not necessarily the same as those obtained by calculating the true average exposures between assumed ages x and $x+1$. If, for example, all the new entrants in an exact age method came under observation three months before their birthdays, by the latest method they would all be treated as entrants at age $x+1$, whereas the true average exposure at age x is $\frac{1}{4}$. More generally, it can be seen that the average exposure at assumed age x for new entrants will be the same by both methods only if k, as well as being the average period between the date of attaining assumed age x and the date of entry, is the proportion of n_x whose date of entry is nearer to the date of attaining assumed age $x+1$ than to the date of attaining assumed age x, e.g. in an exact age method, the proportion whose nearest birthday at entry is $x+1$. The difference in the results by the two methods will seldom be of any significance.

Occasionally, the data may be given in a form suitable for applying the method above, even although the exact dates of the movements are unknown. For example, the movements may be already classified according to the nearest age on the nearest 1st January, although it is of course necessary that the deaths should be grouped according to the nearest age on 1st January preceding death. It is more usual to find the data wholly or partially in this form in examination questions than in actual practice. It should be noted that when the actual dates of the movements are unknown, more accurate results should be obtainable if the data are available in groups according to this method than if all movements, including deaths, are grouped according to age last birthday (or according to the assumed age on 1st January or the policy anniversary before the movement), for, while in the latter case we must assume an even distribution of movements over the year (or assumed year) of age, in the former no such assumption is necessary. Example 2 of Chapter VII illustrates the new method of treating fractional exposures.

CHAPTER VI

PRACTICAL APPLICATION OF THE METHODS

1. Summary of operations.

The methods described in the last three chapters are in general use for carrying out the mortality investigations of life offices, friendly societies and similar institutions, whose records, in the form of policy registers, membership books, valuation cards, etc. provide the necessary information. If we assume that all preliminary questions regarding the limits and scope of the experience, the period to be covered and the objects of the investigation have been decided, we can summarize the various stages in the work thus:

(i) Decide which method of grouping the data is to be used and within that grouping which particular approximations to age are to be adopted.

In making these decisions, we must bear in mind the form in which the information is available and strive to preserve a balance between the demands of accuracy and the limits of practical convenience. We shall generally find, for example, that an exact age method must be rejected either because the necessary information is not available or because the labour is prohibitive.

All life offices and friendly societies carry out every few years, if not annually, valuations of the policies on their books, and for this purpose it is convenient to keep records on a special set of cards. On these are set out the required particulars, which must include some approximation to the ages of the lives on the valuation date, e.g. calendar year of birth, age at one valuation date, etc. It is thus easy to find the approximate age of any life—though possibly not the nearest age—on any valuation date and a calendar year method is therefore possible. The valuation date will generally be 1st January and, if so, we require to know only the calendar year in which each movement occurs and not the exact date, with an obvious

saving of work. This method is generally employed when we do not wish to investigate mortality according to duration since entry and is therefore of more value in friendly society work than in life office investigations where questions of selection arise. These we discuss later (Chapter XI) and for the present content ourselves with noticing the obvious suitability of the policy year method in such a case. The method does, however, involve considerably more work, since movements must be assigned to the policy year in which they occur instead of to the calendar year. Not only does this take longer, but the necessary information is not likely to be so conveniently available.

(ii) Record for each life included in the experience the modes of entry and exit and the ages at entry and exit according to the age approximations selected.

(iii) Construct the exposed to risk formula.

(iv) Use the information recorded in (ii) to calculate the numerical values of the various terms (b_x, n_x, w_x, etc.) included in the formula.

(v) If any constants are included in the formula (e.g. k, h and s in formula (2) of Chapter III), estimate the values to be assigned to them. If this is impossible with the data available—as, for example, when only calendar years of entry and withdrawal are known—it is usual to assume a uniform distribution over the calendar year, i.e. $k=h=\frac{1}{2}$.

(vi) Calculate the values of E_x and hence find the rates of mortality for assumed age x.

(vii) Consider the age distribution of the lives of assumed age x and hence estimate the exact age x to which the rate of mortality at assumed age x corresponds. If this is not an integral age, interpolation will be necessary to find the rates for integral ages.

2. Collection of data.

Whether the investigation is based on the experience of a single office or whether a number of offices are contributing, it is of the first importance to see that clear and full instructions are issued to those who actually prepare the data, so that there is uniformity of practice both in what is to be included in the experience and

in what information is to be recorded. The particular points to be dealt with in the instructions will depend on the scope and objects of the investigation. The following are some of the more important questions which may arise:

(a) *What classes of policy are to be included and for which of them are separate returns to be made?*

There will generally be little doubt regarding the principal classes, but many of the smaller classes require special mention, e.g. whole life policies with a reduced premium in the early years, term and convertible term assurances, joint life assurances, contingent and survivorship assurances, children's deferred assurances after the risk has commenced, etc. If separate returns are to be made, the divisions must be clearly stated and rules laid down so that doubtful classes will be assigned to a definite section. For example, if with and without profit contracts are to be dealt with separately, we must adjudicate on such classes as deferred bonus policies, policies with reduced participation in profits, guaranteed bonus policies, etc. The subdivision of data according to class of policy is discussed in Chapter XV.

(b) *Which cases are to be included in the classes selected for investigation?*

The usual practice is to include only cases which have been passed at tabular rates. Specific instructions are required to cover cases passed at tabular rates without medical examination, cases passed with extra premiums for occupation or with temporary or permanent climatic extras, policies on female lives whether subject to sex extras or not, reassurances from other offices, policies subject to level or decreasing debts, etc.

Under this heading too falls to be considered the question of duplicate policies, i.e. whether if several policies have been issued on the same life they are all to be included in the experience and if not how they are to be treated. (See Chapter XIV.)

(c) *What date of entry is to be given?*

The choice lies between the date of the policy and the date of commencement of risk. In the majority of cases these dates are the

same, but where they differ owing to the policy having been dated back it is usual to adopt the earlier date and to neglect the resulting overstatement of the exposed to risk. This point is of more importance when we are investigating rates of mortality according to the period elapsed since entry.

(*d*) *What date of exit is to be given?*

With deaths and surrenders there is little room for doubt and the date of death and the date of cessation of risk respectively will be given whenever these are known. Exits by lapse require more careful consideration, since the date on which the company's risk ends will not generally be the date of the first unpaid premium and regard must be had to the conditions of the policies providing for extension of cover, payment of death benefit under non-forfeiture regulations, etc. In general, the date of exit for the purposes of the investigation will be the date on which the office's liability definitely ceased. It is probable that the office will be advised of all deaths occurring prior to that date and of few or none after it. The exact method adopted is relatively unimportant, so long as it ensures that exposures and deaths are treated in the same way and that we include no periods of exposure for which we are unable to determine the appropriate deaths and conversely.

3. Mortality experience cards.

We must now consider how best the practical work of an investigation can be carried out. The operations described in para. 1 above appear, and indeed are, simple enough, but it must be remembered that in an investigation of any importance the number of cases to be dealt with will run into thousands or even hundreds of thousands. The mechanics of the task therefore call for careful planning if the labour involved is not to become excessive. The office records themselves will seldom be in a form immediately suitable for our purpose and it is unlikely that we shall find all lives of the same age at entry or exit grouped together. It is therefore convenient, once the decisions called for under para. 2 have been made, to prepare a set of cards—one for each case included in the experience—on which the particulars mentioned in (ii) are entered. Cards have the great advantage that they can be readily sorted into any desired order and by these means

groupings and subgroupings of the data can easily be carried out and enumerated. The exact information recorded on the cards will depend on the method and the approximations to be used, but in all cases the completed cards must supply the assumed ages at entry and exit and the modes of entry and exit.

It will simplify the work if the cards are designed to record the minimum amount of information in the clearest possible way. On the other hand, it is inadvisable to record only the ages and modes of entry and exit, for the possibility of having to recast the method must be borne in mind and much extra work would result if we had to refer back to the various institutions. For an exact age method the following information should be recorded on the cards in addition to such particulars as are necessary to identify the office of origin and the individual life.

Date of birth	: :		
Date of coming under observation	: :	Age	
Date of passing out of observation	: :	Age	
Mode of exit			

It is unnecessary to specify whether the life came under observation as a beginner or a new entrant, for this will be apparent from the date. The ages will be inserted according to the process which it has been decided to follow, e.g. age last birthday for all movements or, if we are to proceed on the lines of (V, 7), age nearest birthday for all movements other than deaths.

Again, if the form of the data is such that method IV of Chapter IV would be adopted, we would record

Year of entry Year of exit Mode of exit	
Age at entry Age on 1/1/30 Age at exit	

The age on 1/1/30 would only be inserted for beginners, i.e. those who entered before 1930. This age and the age at exit would be in accordance with the definitions in (IV, 10).

The reader is recommended to study the forms of card used in some of the principal investigations, e.g. the 1863–93 investigation (Chapter XV) and the offices' annuitants experience 1900–20 (Chapter XVII).

In more recent investigations mechanical methods of recording the data by either the Hollerith or Powers-Samas punched card accounting systems have been used. In 1890 the Hollerith system was used for dealing with the United States census returns and in more recent years the data of the 1900–20 offices' annuitants experience were handled by the Powers system.

For the benefit of those who are not familiar with the principles underlying these systems, it may be as well to add that information is recorded on cards by means of punched holes, the position of the holes as determined by the rows and columns of the card giving the necessary facts. Thus, if we allocate a column of the card to describe the type of policy, a hole punched in position 1 of this column might denote a whole life without profits policy, position 2 a whole life with profits policy and so on. All non-numerical information must of course be coded so that it can be expressed as a number or a series of numbers. An automatic punching machine, operated by a key-board, punches the cards rapidly. The installation includes other machines which can sort the cards according to any designation recorded on them and a tabulator which prints, and if necessary totals, the various items punched on the cards. When large numbers of cards are involved, these mechanical methods are many times faster and more accurate than any method of dealing with the cards by hand.

4. Calculation of b_x, n_x, w_x etc.

Let us now assume that the data have been collected and recorded on the cards and that we are ready to pass to the fourth stage described in para. 1. This can be done by sorting the cards as follows:

(i) Sort according to mode of entry. This gives us two main groups, beginners and new entrants.

(ii) Sub-sort each of the main groups according to assumed age at entry and record the numbers, thus obtaining the values of b_x and n_x for each age x.

(iii) Re-sort all cards according to mode of exit. This gives us three groups, i.e. enders, withdrawals and deaths.

(iv) Sub-sort each of the groups in (iii) according to assumed age at exit and record the number, thus obtaining the values for e_x, w_x and θ_x for each age x.

The final steps in the process can most easily be followed if we have a definite exposed to risk formula in mind.

Let us take formula 2 of Chapter III, viz.

$$E_x = E_{x-1} + sb_{x-1} + (1-s)\,b_x - se_{x-1} - (1-s)\,e_x \\ + kn_{x-1} + (1-k)\,n_x - hw_{x-1} - (1-h)\,w_x - \theta_{x-1}.$$

The only terms whose values are still undetermined are the constants k, h and s.

Once these have been fixed, the final stages of the calculation of E_x are sufficiently indicated by the following table, in which $b'_x = sb_{x-1} + (1-s)\,b_x$, etc.

x	b'_x	e'_x	n'_x	w'_x	θ_x	E_x	q_x

E_x is obtained by adding to E_{x-1} the values of b'_x and n'_x and deducting the values of e'_x, w'_x and θ_{x-1}. At the youngest age for which there are data E_{x-1} is of course zero.

CHAPTER VII

EXAMPLES OF THE CONSTRUCTION OF EXPOSED TO RISK FORMULAE

Example 1.

The following case occurred in an investigation covering the years 1930–34 inclusive:

Date of birth: 1st June 1885.
Date of entry: 1st September 1919.
Date of death: 1st March 1933.

Classify the life, according to the assumed ages at which it entered as a beginner and passed out of observation as a death, by each of the following methods and indicate the approximate equivalent of assumed age x in each case.

(1) Exact age method—all movements grouped by age last birthday.

(2) Exact age method—all movements except deaths grouped by age nearest birthday and deaths grouped by age last birthday.

(3) Calendar year method—beginners grouped by age nearest birthday on 1st January 1930 and similarly for enders, other movements by age nearest birthday on the previous 1st January.

(4) Calendar year method—grouping by calendar years of birth and movement.

(5) Calendar year method—grouping by age last birthday at entry and calendar years of entry and movement.

(6) Policy year method—age last birthday at entry and curtate duration for all movements.

(7) Policy year method—age last birthday at entry and nearest duration for all movements except deaths, for which curtate duration is to be used.

The following table shows the different assumed ages and indicates how they are obtained. The student is recommended to work out the ages himself before referring to the table.

	Assumed age at 1st January 1930	Assumed age at beginning of year of death	Approximate equivalent of assumed age
(1)	44	47	x (exact)
(2)	45	47	x (exact)
(3)	45	48	x
(4)	$1930-1885=45$	$1933-1885=48$	$x-\frac{1}{2}$
(5)	$34+1930-1919=45$	$34+1933-1919=48$	x
(6)	$34+10=44$	$34+13=47$	$x+\frac{1}{2}$
(7)	$34+10=44$	$34+13=47$	$x+\frac{1}{2}$

The year of death is the life, calendar or policy year according to the method in use.

Example 2.

The data obtained from a mortality investigation covering the years 1930–34 are grouped thus:

Nature of movement	Symbol	Definition of x
Beginners on 1st Jan. 1930	b_x	Age last birthday on 1st Jan. 1930
Enders on 31st Dec. 1934	e_x	Age last birthday on 31st Dec. 1934
New entrants	n_x	Age last birthday at entry
Withdrawals	w_x	Age last birthday at exit
Deaths	θ_x	Age nearest birthday at death

Draw up an exposed to risk formula and discuss any assumptions made.

The deaths at nearest age x occur between exact ages $x-\frac{1}{2}$ and $x+\frac{1}{2}$, and, for the reasons given in (V, 6), we therefore require the exposed to risk between those ages. For convenience, let us call $x-\frac{1}{2}$ "assumed age x". All movements other than deaths are given under age last birthday at the date of movement, and movements at age x last birthday occur at exact ages varying between x and $x+1$, i.e. at nearest assumed age $x+1$. All movements may therefore be treated as if they occurred at the nearest assumed age. For example, all the lives included in b_x will be

exposed for a full year at assumed age $x+1$ and not exposed at all at earlier ages; while the lives included in e_x will be exposed for a full year at assumed age x and not exposed at all at assumed age $x+1$.

The exposed to risk formula is thus

$$E_x = E_{x-1} + b_{x-1} - e_{x-1} + n_{x-1} - w_{x-1} - \theta_{x-1}; \quad \text{and} \quad q_x = \frac{\theta_x}{E_x}.$$

q_x is the rate of mortality at assumed age x and gives us an approximation to the rate for exact age $x-\frac{1}{2}$.

We must now consider what approximations we have made. The beginners at age x last birthday have been treated as if they came under observation at exact age $x+\frac{1}{2}$ or assumed age $x+1$, so that for beginners who really came under observation between exact ages x and $x+\frac{1}{2}$, E_x has been understated and for beginners between exact ages $x+\frac{1}{2}$ and $x+1$, E_{x+1} has been overstated. Similarly, for beginners between $x-1$ and $x-\frac{1}{2}$, E_{x-1} has been understated and for beginners between $x-\frac{1}{2}$ and x, E_x has been overstated. These errors will have no effect on the value of E_x, if the understatement due to the first group balances the overstatement due to the last. This will be so if the exact ages at 1st January 1930 of the lives aged between $x-\frac{1}{2}$ and $x+\frac{1}{2}$ are uniformly distributed over that year of age (not the lives between ages x and $x+1$). Similarly, for the withdrawals we can show that the method adopted is correct, if we assume that the exact ages at the dates of withdrawal are uniformly distributed over the year of age $x-\frac{1}{2}$ to $x+\frac{1}{2}$ and so on. No other assumptions are involved and therefore, if these are justified, the exposed to risk formula is correct.

Example 3.

The mortality experience of a body of assured lives is being investigated over the period 1932–36. The numbers of beginners on 1st January 1932 and the numbers existing on 31st December 1936 have been tabulated according to nearest ages on those dates; the numbers of new entrants during the period have been tabulated according to nearest ages at the beginning of the calendar year of entry; and the numbers of deaths and withdrawals have been

tabulated according to nearest ages at the beginning of the calendar year of entry plus the curtate duration at death or withdrawal.

On the assumption that no further information is available explain fully how you would obtain aggregate rates of mortality from the above data.

The reference to ages at the beginning of the calendar year may suggest that the method of age grouping by date should be used. This grouping will, however, be rejected, since the information given regarding the deaths involves the policy rather than the calendar year. We can best deal with the deaths if we take our stand at a policy anniversary and notice that all deaths in the succeeding policy year among lives who, on 1st January before the policy anniversary, were aged x nearest birthday, will be grouped together. We therefore define θ_x as the number of deaths who were of assumed age x on the policy anniversary preceding death, where assumed age x is nearest age x on the preceding 1st January. This grouping must be used without adjustment for the corresponding exposed to risk. We must, therefore, trace our exposed by policy years and find E_x by grouping together lives who were of assumed age x at the beginning of any policy year.

The remaining symbols may be defined as follows:

Definition	Symbol	Method of obtaining x
Beginners on 1st Jan. 1932	b_x	Nearest age on 1st Jan. 1932
Enders on 31st Dec. 1936	e_x	Nearest age on 31st Dec. 1936
New entrants	n_x	Nearest age on 1st Jan. before entry
Withdrawals	w_x	Nearest age on 1st Jan. before the beginning of the policy year of exit

The last two types of movement are very easily dealt with since the classification is strictly consistent with the method used for the deaths. The new entrants n_x came under observation at assumed age x and were therefore exposed for a full year at that age. The withdrawals w_x passed out of observation between assumed ages x and $x+1$ and, assuming that on the average they withdrew halfway through the policy year, they should be treated as exposed

for half a year at age x with no exposure at older ages. The beginners b_x were of assumed age x on their policy anniversaries in 1932 and of assumed age $x-1$ on their policy anniversaries in 1931. If, therefore, we assume policy anniversaries to be evenly distributed over the calendar year, they were under observation on the average for six months at assumed age $x-1$ and for a full year at assumed age x. A similar argument applies to the enders and the exposed to risk formula becomes

$$E_x = E_{x-1} + \tfrac{1}{2}(b_x + b_{x+1}) - \tfrac{1}{2}(e_x + e_{x+1}) + n_x - \tfrac{1}{2}(w_{x-1} + w_x) - \theta_{x-1}.$$

The fraction $\dfrac{\theta_x}{E_x}$ then gives the rate of mortality for assumed age x and it only remains to consider to what exact age this rate applies.

The lives of assumed age x on any policy anniversary were aged x nearest birthday on the preceding 1st January so that their mean age on the anniversary would be $x+l$ approximately, where l was the average period since the preceding 1st January. We have no information regarding the value of l beyond our general knowledge that more policies tend to be effected in the later than in the earlier part of the calendar year (assuming that the office year and the calendar year coincide). When dealing with beginners and enders, where a rough approximation was sufficient, we took $l=\frac{1}{2}$. Without further investigation any attempt to estimate l more accurately is mere guesswork, and we shall therefore take it that the rate of mortality for assumed age x applies to exact age $x+\frac{1}{2}$. Rates for exact integral ages must be found by interpolation.

Example 4.

In an investigation covering the five years 1935–39, the data are available in the following form:

(*a*) Beginners on 1st January 1935 and enders on 31st December 1939, grouped according to age last birthday.

(*b*) New entrants according to age next birthday on 1st January before entry.

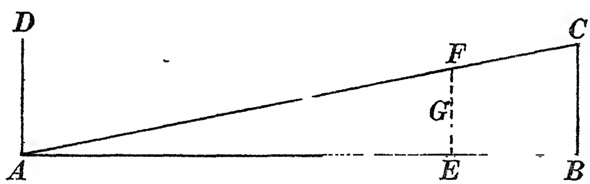
(*c*) Withdrawals and deaths according to age last birthday at date of exit.

No other information being available, how would you calculate rates of mortality?

Note. To this example we give two solutions. In the first, for illustrative purposes, we abandon the principle of (V, 6) and follow what is here the tempting course of altering the grouping of the deaths to fit that of the other movements. In the second solution the principle is followed as hitherto.

First Method. For beginners, enders and new entrants, it would be simplest to define assumed age x as the age last birthday on 1st January and trace the lives through successive calendar years. We should then have to estimate how many deaths occurred in the years 1935–39 among lives aged x last birthday on 1st January before death; in the absence of any details regarding exact dates of death and ages at death we should require to make the usual assumptions of a uniform distribution of deaths over the calendar year and over the year of age. Lives aged x last birthday dying at the beginning of the calendar year will be aged from x to $x+1$ on 1st January before death and those dying at the end of the calendar year will be a year younger on 1st January before death than on the date of death, i.e. will be aged from $x-1$ to x. The lives included in θ_x were therefore aged between $x-1$ and $x+1$ on 1st January before death and, assuming a uniform distribution of deaths, $\frac{1}{2}\theta_x$ would be aged $x-1$ and $\frac{1}{2}\theta_x$ aged x last birthday on that date. As a similar subdivision must be made for θ_{x+1}, the number of deaths among lives aged x last birthday on the previous 1st January will be taken as $\frac{1}{2}(\theta_x+\theta_{x+1})$.

The following diagram shows the position graphically:



The period between the date of death and the previous 1st January is measured along the line AB. At any time t from 1st January, the deaths occurring at age x last birthday may be divided into two groups, a proportion t who were aged $x-1$ last birthday on 1st January and a proportion $1-t$ who were aged x last birthday on 1st January. The line AC is the graph of the number of lives

aged $x-1$ last birthday on 1st January before death and the line DB represents the corresponding number aged x last birthday, for at any point E on the line AB, distant t from A, the ordinates FE and GE are proportional to t and $1-t$ respectively. The sum of these ordinates is constant wherever E is placed on the line AB, and this is consistent with the assumption that deaths are uniformly distributed over the calendar year.

The total number of deaths out of θ_x occurring among lives aged $x-1$ last birthday on 1st January before death is given by the area ACB and for those aged x last birthday by the area ADB. It is evident that these two areas are the same and it follows that $\frac{1}{2}\theta_x$ deaths fall into each of the two groups. A similar subdivision must be made for θ_{x+1} and the number of deaths among lives aged x last birthday on the previous 1st January will be $\frac{1}{2}(\theta_x+\theta_{x+1})$ as before.

The same process may be used for withdrawals, the number aged x last birthday on 1st January before withdrawal being taken as $\frac{1}{2}(w_x+w_{x+1})$. On the average, these lives were exposed for half a year at assumed age x, so that $\frac{1}{4}(w_x+w_{x+1})$ must be deducted from E_{x-1} in obtaining E_x and the same term deducted from E_x in obtaining E_{x+1}. The terms involving w in the formula for E_x will therefore be $\frac{1}{4}(w_{x-1}+2w_x+w_{x+1})$.

Alternatively, we may argue that we are approximating to the rate of mortality for age x last birthday and that the withdrawals included in w_x which take place at age x last birthday should not therefore contribute to the exposed to risk at assumed age x. In strict theory neither method is correct, and as the first expression is complicated and the difference between it and w_x will usually be small—the two are equal if $w_x=\frac{1}{2}(w_{x-1}+w_{x+1})$—we may well prefer to adopt the simpler form. The student who wishes to pursue this matter further is referred to example 5.

The other movements present no difficulty and the table on p. 63 sets out the position.

The last column gives the assumed age for purposes of the exposed to risk formula. Some assumption must be made as to the average length of the period between the date of entry for n_x and the preceding 1st January, and if this be taken as half a year,

Definition	Symbol	Method of obtaining x	Assumed age
Beginners on 1st Jan. 1935	b_x	Age last birthday	x
Enders on 31st Dec. 1939	e_x	Age last birthday	x
New entrants	n_x	Age next birthday on preceding 1st Jan.	$x-1$ on preceding 1st Jan.
Withdrawals	w_x	Age last birthday	x

n_x will be treated as exposed for half a year at assumed age $x-1$ and a full year at assumed age x. The exposed to risk formula is, therefore,

$$E_x = E_{x-1} + b_x - e_x + \tfrac{1}{2}(n_x + n_{x+1}) - w_x - \tfrac{1}{2}(\theta_{x-1} + \theta_x).$$

We then have
$$q_x = \frac{\frac{1}{2}(\theta_x + \theta_{x+1})}{E_x}.$$

q_x is the rate of mortality applicable to lives aged x last birthday on any 1st January from 1935 to 1938, and, assuming as before an even distribution of birthdays over the calendar year, this may be taken as an approximation to the rate for age $x+\frac{1}{2}$.

Second Method. In this method we retain the deaths in their original groups and estimate the exposed to risk at age x last birthday to correspond. The beginners and enders are suitably grouped, but we do not know the average length of the period from the preceding birthday to 1st January 1935 or 31st December 1939. Let us assume that this is six months and that for lives included in w_x the average period between the preceding birthday and the date of withdrawal is also six months.

This leaves n_x, which presents the same problem as w_x in the first method. The lives included in n_x were aged x next birthday on 1st January before entry and, assuming an even distribution of entry dates over the calendar year and of entry ages over the year of age $x-1$ to x, they were on the average aged x at the actual date of entry. As we are calculating the rate of mortality for exact age x, we shall assume that n_x are exposed for a full year at that age and not exposed at all at age $x-1$.

The assumed ages for the various movements are then as follows:

Symbol	Assumed age
b_x	x on preceding birthday
e_x	x on preceding birthday
n_x	x
w_x	x on preceding birthday
θ_x	x on preceding birthday

The exposed to risk formula is therefore:

$$E_x = E_{x-1} + \tfrac{1}{2}(b_{x-1} + b_x) - \tfrac{1}{2}(e_{x-1} + e_x) + n_x - \tfrac{1}{2}(w_{x-1} + w_x) - \theta_{x-1};$$

and

$$q_x = \frac{\theta_x}{E_x}.$$

The working rule given in (V, 6) indicates that the second method is the better, although it involves more approximations in the formula for E_x. The example may be used to illustrate the statement in (V, 6) that a considerable part of E_x is unaffected by errors due to approximations to the length of exposure at the age of entry or exit. Consider, for example, the beginners b_{x-1} in the second formula. We include $\tfrac{1}{2}b_{x-1}$ in E_{x-1} and another $\tfrac{1}{2}b_{x-1}$, making b_{x-1} in all, in E_x. The assumption that b_{x-1} are exposed for half a year at age $x-1$ may be incorrect, but it is undoubtedly correct to treat them as exposed for a full year at age x, since lives aged $x-1$ last birthday on 1st January 1935 cannot be more than age x on that date. The same argument applies to e and w and we see therefore that the greater part of E_x is accurate, the only possibility of error being in the terms $\tfrac{1}{2}b_x$, $\tfrac{1}{2}e_x$, $\tfrac{1}{2}w_x$ and n_x. (It will be remembered that n_x is an approximation to the number of entrants at exact age x.) Clearly therefore better results should be obtained by the second than by the first method, which involves an approximation to the number of deaths at assumed age x, so that the whole numerator is subject to the risk of error.

Example 5.

Note. This example does not introduce any new principles and is intended for the benefit of any readers who are not fully satisfied with the approximations used in the first method of example 4 in obtaining

the terms in w. It may be regarded as a fairly difficult exercise in manipulating the functions used in exposed to risk formulae.

The exposed to risk for lives aged x last birthday on 1st January is required and the withdrawals are grouped according to the age last birthday at withdrawal.

Half the withdrawals at age x last birthday were taken to be aged x last birthday and the other half aged $x-1$ last birthday on 1st January before withdrawal; the total number aged x last birthday on 1st January before withdrawal was then $\frac{1}{2}(w_x+w_{x+1})$. This is quite correct on the assumption of an even distribution of withdrawals over the years of age x to $x+1$, etc. and over the calendar year. The next step was to say that these $\frac{1}{2}(w_x+w_{x+1})$ lives would, on the same assumptions, be exposed for half a year on the average between assumed ages x and $x+1$. This is not in general correct, for, unless $w_x=w_{x+1}$, the dates of withdrawal will not be uniformly spread over the calendar year.

Consider the diagram in Example 4 which, although it refers to deaths, is equally applicable to withdrawals. If t is the fractional period between the preceding 1st January and the date of withdrawal, represented by AE in the diagram, the ordinate GE is proportional to $1-t$ and represents the number of lives who withdraw at duration t from the beginning of the calendar year of withdrawal, who were aged x last birthday at the date of withdrawal and who were also aged x last birthday on the preceding 1st January. The exposure which should be allowed for each of these lives at assumed age x is t and the average exposure for all the $\frac{1}{2}w_x$ withdrawals answering this description when $0 \leqslant t \leqslant 1$ is

$$\frac{\int_0^1 (1-t)\,t\,dt}{\int_0^1 (1-t)\,dt}.$$

The expression reduces to $\frac{1}{3}$, whereas previously we took this average exposure as $\frac{1}{2}$.

As regards the $\frac{1}{2}w_{x+1}$ lives who were aged $x+1$ last birthday at withdrawal and x last birthday on the preceding 1st January, the

number withdrawing at duration t from 1st January is proportional to t and the average exposure is

$$\frac{\int_0^1 t^2\,dt}{\int_0^1 t\,dt} = \frac{2}{3}.$$

$\frac{1}{2}w_x$ should therefore be exposed for $\frac{1}{3}$ of a year and $\frac{1}{2}w_{x+1}$ for $\frac{2}{3}$ of a year at assumed age x, so that the appropriate deduction in the formula for E_x is $\frac{1}{3}w_x + \frac{1}{6}w_{x+1}$. Similarly, $\frac{1}{3}w_{x-1} + \frac{1}{6}w_x$ should be deducted in the formula for E_{x-1} and the remainder of $\frac{1}{2}(w_{x-1} + w_x)$, i.e. $\frac{1}{6}w_{x-1} + \frac{1}{3}w_x$, must be deducted in the formula for E_x. The terms involving w in the latter formula should therefore be $\frac{1}{6}w_{x-1} + \frac{2}{3}w_x + \frac{1}{6}w_{x+1}$ instead of $\frac{1}{4}w_{x-1} + \frac{1}{2}w_x + \frac{1}{4}w_{x+1}$.

The alternative method assumed in effect that w_x could be taken as a close approximation to the number of withdrawals aged x last birthday on 1st January nearest to the date of withdrawal, for on this assumption all the lives in w_x may be treated as withdrawing at assumed age x. The following argument might be advanced to justify this assumption. The lives withdrawing at duration t from 1st January ($0 \leqslant t \leqslant \frac{1}{2}$), who were aged x last birthday at the date of withdrawal, would be aged between $x-t$ and $x-t+1$ on the nearest 1st January. A proportion t would therefore be aged $x-1$ last birthday and a proportion $1-t$ aged x last birthday on that date. The total number aged $x-1$ last birthday and x last birthday for all values of t between 0 and $\frac{1}{2}$ would therefore be

$$w_x \int_0^{\frac{1}{2}} t\,dt \quad \text{and} \quad w_x \int_0^{\frac{1}{2}} (1-t)\,dt.$$

These expressions reduce to $\frac{1}{8}w_x$ and $\frac{3}{8}w_x$ respectively.

Similarly, dealing with the $\frac{1}{2}w_x$ withdrawals in the latter half of the calendar year, $\frac{3}{8}w_x$ would be aged x last birthday on the nearest 1st January to withdrawal and $\frac{1}{8}w_x$ aged $x+1$ last birthday. The number of withdrawals aged x last birthday on the nearest 1st January to the date of withdrawal would therefore be $\frac{1}{8}w_{x-1} + \frac{3}{4}w_x + \frac{1}{8}w_{x+1}$, and this would be the term in w in the formula for E_x. w_x is quite a close approximation unless the

values of w_{x-1}, w_x and w_{x+1} are far removed from an arithmetic progression.

There is, however, a fallacy in this argument which accounts for the difference between the resulting expression and that obtained previously. Consider the withdrawals aged x last birthday at the beginning of the year of withdrawal, i.e. $\frac{1}{2}(w_x+w_{x+1})$. E_x has been understated by treating the $\frac{3}{8}w_x+\frac{1}{8}w_{x+1}$ withdrawals in the first six months of the calendar year as if they withdrew at assumed age x, i.e. as if withdrawal occurred on 1st January instead of in the succeeding six months, and E_x has been overstated by treating the $\frac{1}{8}w_x+\frac{3}{8}w_{x+1}$ withdrawals in the second six months as if they withdrew at assumed age $x+1$. In effect, we have assumed that these understatements and overstatements will balance, but this will not be the case unless $w_x=w_{x+1}$. An attempt to correct for the resulting error will lead to the same argument as was used in adjusting for the error introduced by the first method, and we shall be left with $\frac{1}{6}w_{x-1}+\frac{2}{3}w_x+\frac{1}{6}w_{x+1}$ as the correct term involving w in the exposed to risk formula on the stated assumptions regarding the distribution of withdrawals.

In general, either of the forms used in Example 4 would be sufficiently accurate for practical purposes and, as already indicated, the full investigation has been made mainly as an exercise. It serves however to show the need for careful reasoning in handling the functions used in exposed to risk formulae if inconsistencies are to be avoided. Even if there is no special virtue in the assumptions made as to the distribution of the movements over the year of age and the calendar or policy year, it is important that the operator should not unknowingly make other and inconsistent assumptions during the course of the work. It is of course a different matter if the operator recognizes that a particular procedure involves inconsistencies and decides that, for practical purposes, this is of no importance.

Examination questions.

The majority of readers will, no doubt, be students and, although it is not intended to usurp the functions of a tutor, some remarks on the subject of examinations may be helpful. Examination questions

on exposed to risk formulae are not usually of the kind for which adequate preparation can be made by "cramming", and students should abandon any idea that, by working or studying a sufficient number of examples, they will be able to have ready-made answers to all possible questions. Rules may be devised which will provide a partial check on exposed to risk formulae, but the use of such rules except as a check is most undesirable. In any case, the student who wishes to become an actuary as opposed to a person who has merely passed certain examinations will not welcome any short cut which does not require a proper study and complete understanding of principles. The examples which have been given in this chapter do not therefore attempt to exhaust the many and varied types of question which may be set. They may, however, help to show the general lines of approach which may be followed and to illustrate some of the principles which have been laid down.

Finally, a few hints on the way to set out answers to examination questions on exposed to risk formulae are perhaps permissible. Actuarial terminology is not by any means standardized and it is therefore of primary importance that ambiguous phrases should be avoided, especially when giving definitions. At an early stage in the answer the meaning to be assigned to E_x should be explained and all the symbols to be used clearly defined. If the method of age grouping for any type of movement is inconsistent with the definition of E_x, the approximations necessary to bring the former into line with the latter must be explained. This arises when, as in the examples in this chapter, the age grouping is laid down in the question and it is stated that no other information is available. If, on the other hand, full details as to dates of birth, dates of movement, etc., are said to be available, the answer should state what method or methods of age grouping would be used and should give reasons for the choice. It is not sufficient merely to state that a life, calendar or policy year method is suitable.

CHAPTER VIII

ERRORS OF APPROXIMATION

1. No one studying the subject of exposed to risk formulae can fail to remark how frequently the demands of practice cause us to modify the requirements of exact theory. In the course of our analysis we have used numerous approximations regarding distributions of birthdays and movements, as, for example, average ages at entry and exit, average durations at withdrawal, uniform distributions of birthdays over the calendar year, etc. Most of these are natural and reasonable and it seems obvious that they cannot lead us into serious error. Moreover, we can investigate the limits of their probable accuracy by suitably sampling the data.

On several occasions however—either to simplify the work or to overcome deficiencies in the data—we have made assumptions of a rather different character, where neither the exact nature of the assumption nor the error thereby introduced are so readily apparent.

In this chapter we shall consider a number of assumptions and errors of this type.

2. Uniform distribution of deaths.

In (II, 5), we saw that our method of treating fractional exposures involved the assumption that, among a group of lives under observation from age x to age $x+1$ or previous death, the deaths between ages x and $x+t$ (where $0 \leqslant t \leqslant 1$) would be a fraction t of the total deaths between ages x and $x+1$. For this to be true the data must clearly be of sufficient extent to avoid large random errors, but, this being granted, let us see how closely the assumption fits the facts.

Consider the distribution of deaths between ages x and $x+1$ in a large group of P_x lives aged x which is affected by no movement other than death. Then

$$\theta_x = \int_0^1 P_{x+t}\mu_{x+t}\,dt,$$

the deaths occurring in the infinitesimal interval of time dt between ages $x+t$ and $x+t+dt$ being given by $P_{x+t}\mu_{x+t}dt$. We can therefore estimate the accuracy of our assumption in a particular case by examining the progression of $P_x\mu_x$ and more generally by examining the progression of $l_x\mu_x$ in a standard mortality table. (We confine the use of the symbols l_x and d_x to the number of lives aged x and the number of deaths between ages x and $x+1$ according to a life table, where the relation $l_{x+1}=l_x-d_x$ holds. The functions P_x, P_{x+1} and θ_x are not connected by a similar relation.) In the A 1924–29 Table $l_x\mu_x$ increases slowly from age 30 up to about age 75 and decreases thereafter. The value of $\frac{l_{x+1}\mu_{x+1}}{l_x\mu_x}$ never differs from unity by more than 10 per cent. If therefore we were to base our estimate of q_x solely on observations covering a fraction of the year of age x to $x+1$, we might thereby introduce an error not exceeding and probably considerably less than 10 per cent.

In practice, the resulting error in q_x will be very much less, since fractional exposures usually constitute only a small proportion of the whole. Moreover, some of the fractional exposures (enders and withdrawals) relate to ages x to $x+t$ and others (beginners and new entrants) to ages $x+s$ to $x+1$. The errors introduced will therefore tend to cancel each other and the final effect on q_x will usually be negligible.

Special cases may however arise where care will be required. Suppose, for example, that we are investigating an endowment assurance experience in which a large number of policies mature between ages 64 and 65. The exposed to risk at age 64 will then consist largely of lives who were under observation only in the earlier part of the year of age. Again, at a young age, say 20, the greater part of the value of E_{20} might lie in the term $(1-k)\,n_{20}$ and E_{20} would therefore relate principally to the second half of the year of age. In such cases a special investigation into the distribution of the exposed to risk might have to be made and some adjustment might be necessary in the value of q_x. Actually, in the second example, $l_x\mu_x$ would normally be so nearly constant in the region of age 20 as to make the error negligible.

3. Substitution of b_x and e_x for P_x^{29} and P_x^{34}.

Formula (4) of Chapter II will give the same result as formula (3) provided that

$$P_x^{29}-b_x=P_x^{34}-e_x,$$

i.e. provided that P_x^{29} and P_x^{34} give rise to the same number of deaths before the end of the years 1929 and 1934 respectively, assuming that there are no new entrants or withdrawals. The error introduced by using formula (4) is usually insignificant.

As an example, suppose that the functions have the following values:

$$P_x^{29}=10, \qquad P_x^{34}=2000,$$
$$P_x' =3100, \qquad \theta_x =500.$$

Then, if we take s to be $\frac{1}{2}$, formula (3) gives

$$q_x=\frac{500}{5+3100+1000}=\cdot 1218.$$

Now suppose that $b_x=10$ and $e_x=1900$. Then, by formula (4),

$$q_x=\frac{500}{5+3100+2000-950}=\cdot 1203.$$

In spite of the fact that an extreme case has been taken, the error is only about 1 per cent of q_x, and a closer approximation should result if q_x is small or if the disparity between P_x^{29} and P_x^{34} is not so great.

The general argument remains unchanged when there are new entrants and withdrawals, except that the proviso at the beginning of the paragraph will require the total reduction in the number of lives under observation to be the same in each case, i.e. the number of deaths plus the number of withdrawals less the number of new entrants.

4. Treatment of fractional exposures.

Consider a group of lives whose rate of mortality at exact age x is q_x. Suppose that a number of other lives experiencing the same rate of mortality join the group between ages x and $x+1$ at average age $x+k$. In investigating the mortality of the group, it is reasonable that we should fix the exposure for the additional lives in such a way that the resulting rate of mortality will still be q_x. Let n_x be the number of the additional new entrants. The deaths among the

new entrants will number $n_x \times \dfrac{l_{x+k} - l_{x+1}}{l_{x+k}}$ and, assuming a uniform distribution of deaths, this reduces to $\dfrac{n_x q_x (1-k)}{1-kq_x}$. We should therefore include the term $\dfrac{(1-k)n_x}{1-kq_x}$ in E_x instead of $(1-k)n_x$.

The reader will appreciate that, in actual fact, the rate of mortality for the new entrants would probably differ from that for the original group on account of random errors due to paucity of data. What we have done in this paragraph is to frame the term containing n_x in such a way that, in the ideal case where the data are unlimited and there are therefore no random errors, the rate of mortality for the original group would be unaltered by the inclusion of the new entrants. The fact that the ideal is never attained does not vitiate the argument. The main object of including fractional periods of exposure is, of course, to increase the amount of data with a view to reducing the size of the random error in the crude value of q_x. The argument of this paragraph applies to all types of fractional exposures.

As regards withdrawals, it will be remembered that in (II, 6) we postulated that lives dying between ages x and $x+1$ should be included in E_x to the same extent as if they had not died and we pointed out in (II, 10) that this rule was not strictly observed in formula (4) of that chapter. The correctness of the rule may not be readily apparent and an example may help to demonstrate it. Consider a group of 100 lives aged x exactly, of whom 40 die before age $x+1$, the rate of mortality being therefore ·4. Now suppose that half the survivors to exact age $x+\frac{1}{2}$ withdraw from the experience at that age. Assuming an even distribution of deaths, the survivors will number 80, of whom 40 will withdraw; the number of deaths among the remainder in the second half of the year of life will be proportionately reduced to 10. If we give a full year's exposure to all the deaths, the rate of mortality is $\dfrac{20+10}{100-\frac{1}{2}40} = \cdot 375$. If, on the other hand, we apply the rule laid down in (II, 6), we must argue that half the lives dying in the first half of the year of life would have withdrawn at age $x+\frac{1}{2}$ if

they had not previously died and that the exposed to risk should therefore be taken to be $100-\frac{1}{2}40-\frac{1}{2}10=75$, leading to the correct rate of mortality, $\cdot 4$.

More generally, we should deduct a fraction $1-h$ not only for each of the w_x withdrawals, but for each of the $\dfrac{w_x}{1-hq_x}$ lives of exact age x required to provide w_x survivors.

If desired, we can approach the problem of the withdrawals in a similar way to that adopted for the new entrants. Consider a group of lives of exact age x under review until age $x+1$ or previous death, experiencing rate of mortality q_x. Now suppose that w_x withdraw at age $x+h$ on the average. Assuming that these lives experience the same rate of mortality as the group as a whole, the number of deaths excluded from the experience will be $w_x \times \dfrac{l_{x+h}-l_{x+1}}{l_{x+h}}$, which on the usual assumption reduces to $\dfrac{w_x q_x(1-h)}{1-hq_x}$. A deduction of $\dfrac{w_x(1-h)}{1-hq_x}$ must therefore be made from the previous value of E_x in order that the rate of mortality will be unchanged. This expression is the same as that brought out above.

A similar approach for beginners and enders leads to the following equation for the rate of mortality.

$$\theta_x=q_x\left(P_x+\frac{(1-s)\,b_x}{1-sq_x}-\frac{(1-s)\,e_x}{1-sq_x}+\frac{(1-k)\,n_x}{1-kq_x}-\frac{(1-h)\,w_x}{1-hq_x}\right).$$

This formula is unnecessarily complicated for practical use. Furthermore it is difficult to justify such refinements when the arguments all depend on an assumption which is only approximately correct, i.e. the assumption that deaths are uniformly spread over the year of age. The error introduced by using formula (5) of Chapter II is unlikely to be of practical importance unless q_x is very large and it is when q_x is large that the correctness of the assumption is most open to doubt.

5. Constancy of *k*, *h* and *s*.

The practical value of exposed to risk formulae depends on the assumption that k, h and s are approximately constant at all ages or at least over a considerable range of ages. On general grounds

this appears reasonable enough. Supposing for the moment that we are dealing with the ideal case of unlimited data, so that the true values of the fractions are obtained, it is difficult to imagine circumstances which would cause a large variation.

Differences will however occur owing to insufficiency of data. Errors arising in this way are statistical and some idea of their size can easily be obtained by statistical methods. For example, if we assume that withdrawals between ages x and $x+1$ are equally likely to occur at any intermediate point of age, then the true mean exposure at age x for lives withdrawing at that age will be $\frac{1}{2}$. For any particular group of w_x withdrawals the mean exposure will differ from the true mean and the greater the value of w_x, the smaller the difference is likely to be. Statistical theory shows that this difference is most unlikely to exceed $\frac{1}{\sqrt{w_x}}$ (i.e. approximately three times the standard error of the mean). Thus if there are 100 withdrawals at age x, the error will almost certainly be less than ·1, involving an error of 10 in E_x, which will not normally be of any practical importance, particularly when it is remembered that statistical errors in q_x should largely be removed by satisfactory graduation. When w_x is very small, the error in h may approach its limit of ·5, but as w_x will then probably be small relative to E_x the error will still be insignificant.

6. Use of exact age $x+l$.

In several of our methods we arrived at the rate of mortality for an assumed age which we took to be the rate of mortality for some exact age $x+l$. To examine the error thereby introduced, let us consider the methods of (IV, 10) and (IV, 11) where the greatest departure from strict accuracy occurred. The exposed to risk is here spread over three years of age (e.g. $x-1$ to $x+2$) instead of one and we assume that this gives us the rate of mortality applicable to age x. Here again, paucity of data may lead to an uneven distribution of the exposed to risk over the three years of age, but even when E_x is as small as 100 the error in the equivalent exact age is unlikely to exceed ·25 of a year of age, which should not cause an error of more than $2\frac{1}{2}$ per cent of q_x even before graduation.

Even if the exposed to risk is uniformly distributed over the years of age $x-1$ to $x+2$, it is not correct to take x as the exact age to which the resulting rate of mortality applies, since second differences of q_x are not necessarily negligible. The distribution of the exposed to risk is (IV, 12)

$\frac{1}{6}$ between ages $x-1$ and x,
$\frac{2}{3}$,, ,, x and $x+1$,
$\frac{1}{6}$,, ,, $x+1$ and $x+2$,

and a very rough idea of the error can be obtained by comparing $\frac{1}{6}q_{x-1}+\frac{2}{3}q_x+\frac{1}{6}q_{x+1}$ with q_x for specimen ages according to standard tables. We find that the approximation nearly always results in an overstatement of the true value of q_x, but seldom by more than 2 per mille. This error is even smaller than those already mentioned, but it is important to notice that its effect is slightly to displace the whole q_x curve. An error of this kind which always operates in the same direction is more serious than one of the same average size whose direction varies from age to age, since the former will not be reduced by graduation. Even so, it will not be of practical significance.

7. Entry and exit in the same year of age.

In (II, 9) we assumed that the average ages at entry and exit of new entrants and withdrawals between ages x and $x+1$ were $x+k$ and $x+h$ respectively. Some lives would probably be included in both n_x and w_x and the exposure at age x for each of them would be taken as $h-k$, which if $k>h$ would be negative. At first sight this may seem absurd, but the fact is that, in taking $x+h$ as the average age at withdrawal, we are dealing with the withdrawals between ages x and $x+1$ as a whole and it is unsound to apply the result to a particular case or to a particular section of the withdrawals. The fraction h we presumed to have been calculated from a representative sample of the withdrawals which would include cases where entry and withdrawal occurred in the same year of age.

Sometimes, however, the values of the constants must be fixed by analytical rather than statistical methods—by argument instead

of by calculation. The reasoning behind such estimates is necessarily loose. We may argue, for example, that new entrants in an experience of assured lives have good reason to effect policies shortly before a birthday, so that the period between the date of entry and the next birthday is likely to be less than half a year. On the other hand, it may be argued that there is no reason to expect any such factor affecting the age distribution of the withdrawals and that an even distribution over the year of age is therefore likely, in which case the withdrawals should be exposed to risk for half a year on the average. It should be noted that this argument neglects the possibility of withdrawals occurring mainly at or near the policy anniversary, a feature which might affect the age distribution.

It may be worth considering exactly what we are assuming when we determine the values of the constants by general reasoning. It is clear that the distribution of the withdrawals, for example, is to some extent dependent on the number and age distribution of beginners, enders, new entrants and deaths at the age under review, for these factors affect the incidence of the withdrawals. For instance, a large number of new entrants will lend weight to the later part of the year of age at which they enter and, other things being equal, a larger proportion of the withdrawals will occur in the later part of the year of age than if there were comparatively few new entrants. By allowing an average of half a year's exposure at the age of withdrawal, we imply that an even distribution of withdrawals holds for that particular experience; it would not necessarily hold if there were a change in the number or distribution of the other movements.

This is a convenient but rather unconvincing assumption and it appears more reasonable to assume that an even distribution of withdrawals would occur if there were no other movements—an assumption which could be consistently applied no matter how the numbers of the other movements varied. Let us assume then that the new entrants are uniformly distributed over the year of age x to $x+1$ and that the withdrawals would be uniformly distributed if there were no new entrants. Neglecting the effect of any other movements, we shall try to estimate the average exposure for lives who both enter and withdraw between x and $x+1$.

If μ^w_{x+t} is the force of withdrawal at age $x+t$, where $0 \leqslant t \leqslant 1$, and P^w_{x+t} the survivors of P^w_x lives aged x exactly when withdrawal is the only cause of decrement and there are no causes of increment, $\mu^w_{x+t} P^w_{x+t}$ is constant. This is merely another way of saying that the withdrawals are evenly spread over the year of age. Unless the rate of withdrawal is high, P^w_{x+t} will not decrease much as t increases from 0 to 1 and we can assume that the proportion of those aged $x+t$ withdrawing before attaining $x+1$ is $(1-t)N$, where N is constant for all values of t between 0 and 1. Let n_x be the number of new entrants between ages x and $x+1$, so that $n_x \Delta t$ will be the number entering between ages $x+t$ and $x+t+\Delta t$. The number of these withdrawing before attaining age $x+1$ will therefore be $N(1-t)\, n_x \Delta t$, provided that Δt is small, and the total number withdrawing out of the n_x new entrants will be

$$Nn_x \int_0^1 (1-t)\, dt = \frac{Nn_x}{2}.$$

The lives who enter at age $x+t$ and withdraw before attaining age $x+1$ will on the average be exposed for a fraction $\dfrac{1-t}{2}$ of a year of age, and the total exposure will therefore be

$$Nn_x \int_0^1 \frac{(1-t)^2}{2}\, dt = \frac{Nn_x}{6}.$$

The average exposure is therefore one-third of a year of age. In this case our assumption for practical purposes would be

$$h = k = \tfrac{1}{2},$$

and therefore entrants and withdrawals in the same year would be credited with no exposure. For such cases the exposure would be underestimated by one-third of a year. When it is remembered that the new entrants will normally be only a small proportion of the total exposures and that only a small proportion of the new entrants will withdraw in the year of entry, it will be seen that only in very exceptional cases could this error be of practical significance. Similar arguments can be applied to other cases of lives coming under and passing out of observation in the same year of age, e.g. beginners withdrawing before their birthdays in the first year of the experience.

8. Our examination of the errors introduced by the various approximations we have made has led us to the negative but gratifying conclusion that under normal conditions these errors can safely be neglected. At the same time it has been made apparent that the price of accuracy is unceasing vigilance; the investigator must be constantly on the alert for the exceptional case where the normal methods and assumptions do not apply and where special treatment is demanded. It is a good rule to suspect and to enquire carefully into every feature of the data which appears in the slightest degree abnormal and not to assume without full investigation that its effect on the resulting rates of mortality can be ignored.

CHAPTER IX

THE CENTRAL RATE OF MORTALITY

1. Definition of the central rate of mortality.

We define m_x, the *central rate of mortality* experienced by a group of lives at age x, as the ratio of the number of deaths in the group between ages x and $x+1$ to the number of years of exposure to the risk of death, there being no causes of increment or decrement other than death and *each death being treated as exposed only up to the age of death.*

This differs from the definition of q_x in (II, 11) only in the treatment of the deaths. In arriving at q_x these are exposed for a full year of age instead of being exposed from age x up to the age of death only.

The definition is applicable whether it refers to exact age x or to one of the variations which may be grouped under the description "assumed age x".

2. Exposed to risk formulae.

It is quite a simple process to devise these formulae in what we may call central form, i.e. formulae which will lead directly to m_x. We denote the *central exposed to risk* at assumed age x by E_x^c, so that $m_x = \dfrac{\theta_x}{E_x^c}$. The formulae for E_x provide a convenient starting point, since the only difference between E_x and E_x^c is in the treatment of the deaths.

If we assume that the deaths are uniformly distributed over the year of age, they should on the average be treated as exposed for half a year in E_x^c as compared with a whole year in E_x. Hence

$$\left.\begin{aligned} E_x^c &= E_x - \tfrac{1}{2}\theta_x, \\ E_{x-1}^c &= E_{x-1} - \tfrac{1}{2}\theta_{x-1} \end{aligned}\right\}, \qquad \ldots\ldots(1)$$

from which we obtain the relation

$$E_x^c = E_{x-1}^c + (E_x - E_{x-1}) + \tfrac{1}{2}(\theta_{x-1} - \theta_x).$$

If our exposed to risk formula for E_x is

$$E_x = E_{x-1} + \phi_x - \theta_{x-1},$$

where ϕ_x consists of the terms involving b, e, n and w, we have

$$E_x^c = E_{x-1}^c + \phi_x - \tfrac{1}{2}(\theta_{x-1} + \theta_x), \qquad \ldots\ldots(2)$$

irrespective of the form of ϕ_x. This is our central exposed to risk formula.

Formula (1) leads at once to the well-known approximate relation

$$m_x = \frac{2q_x}{2 - q_x}, \qquad \ldots\ldots(3)$$

since $$m_x = \frac{\theta_x}{E_x^c} = \frac{\theta_x}{E_x - \tfrac{1}{2}\theta_x}.$$

3. Errors introduced by central exposed to risk formulae.

It is clear from the definition of m_x that the distribution of deaths has a greater effect on m_x than on q_x, and a more serious error is accordingly introduced if an even distribution is wrongly assumed. If the average age at death is less than $x+\frac{1}{2}$, the total exposure in E_x^c attributable to lives dying between x and $x+1$ will be overstated and m_x will be understated. The reverse will be true if the average age at death is more than $x+\frac{1}{2}$. It was explained in (VIII, 2) that the distribution of the deaths in a life table depends on the variation of the function $\mu_x l_x$ over the year of age. Even where this function is changing rapidly and the distribution of deaths is therefore far from uniform, the error in m_x may not be of practical importance, for if m_x is small and, in consequence, $\frac{1}{2}\theta_x$ is small in relation to E_x^c, even a comparatively large error in the total exposure for the lives dying at age x may amount to a negligible proportion of E_x^c.

In the A 1924–29 table, for example, where $\mu_x l_x$ is increasing rapidly at age 55, an average exposure of half a year for the deaths at that age involves an overstatement in m_{55} of only about 1 in 10,000. At age 90, on the other hand, where $\mu_x l_x$ is falling rapidly, there is an understatement in m_{90} of about $\frac{1}{2}$ per cent. Errors of this kind may operate in the same direction over a considerable range of ages and it may be unsafe to ignore them even if they do not exceed $\frac{1}{2}$ per cent.

It should be noticed that we do not eliminate the error in m_x due to the assumption of a uniform distribution of deaths by first calculating q_x and then obtaining m_x by relation (3), since the assumption is inherent in this relation.

It is of interest to note also that the assumption as to the uniform distribution of deaths made in para. 2 relates to the actual deaths experienced. This is not the same assumption as that made in earlier chapters, i.e. that the distribution would be uniform if there were no other movements. The difference is of no practical importance, but for the sake of completeness we shall discuss the matter in some detail in para. 6.

4. Alternative formula for E_x^c.

We can approach the problem of calculating E_x^c in quite a different way. In the conditions laid down in the definition of m_x, let P'_x be the number of lives in the group aged x and P'_{x+t} the number of survivors to age $x+t$.

If the year of age is divided into n parts where n is large, the lives in the group will contribute $\frac{1}{n}P'_{x+\frac{s}{n}}$ to E_x^c between ages $x+\frac{s}{n}$ and $x+\frac{s+1}{n}$. Hence, we have

$$E_x^c = \sum_{s=0}^{n-1} \frac{1}{n} P'_{x+\frac{s}{n}}.$$

When n tends to infinity, this relation takes the form

$$E_x^c = \int_0^1 P'_{x+t}\,dt. \qquad \ldots\ldots(4)$$

The reader will see that there is a difference between this method of deriving the exposed to risk and those used previously. Instead of considering the exposed to risk as the sum of the contributions made by individual lives, we regard it as the sum of the contributions made by the groups of lives in the different intervals of age x to $x+\frac{1}{n}$, etc.

If we assume that the deaths are uniformly spread over the year of age, formula (4) becomes

$$E_x^c = \int_0^1 (P'_x - t\theta_x)\, dt = P'_x - \tfrac{1}{2}\theta_x.$$

Since in the above conditions $q_x = \dfrac{\theta_x}{P'_x}$, we obtain once again formula (3).

5. Alternative formula for E_x^c: fractional exposures.

For lives entering the experience at age $x+s$ and leaving it at age $x+s+k$ $(0 \leqslant s < s+k \leqslant 1)$ we shall include in E_x^c the appropriate fractional exposure k. For the majority of lives, either s will be zero or $s+k$ will be unity, or both.

Let P_{x+t} be the number of lives attaining age $x+t$ during the investigation, whether or not they were under observation at age x and continued under observation until age $x+1$ or previous death. Then the argument of para. 4 can be repeated and we have

$$E_x^c = \int_0^1 P_{x+t}\, dt. \qquad \text{......(5)}$$

Formula (4) is a special case of formula (5).

The new formula is not of much practical value as it stands, for it is usually impossible to devise a simple expression for P_{x+t} which will lend itself to integration. Its chief virtue lies in the fact that it leads to the census method with which we shall deal in the next chapter.

Once again, x need not necessarily be exact age x, but may be assumed age x. If, for example, this is defined as nearest age x on 1st January in any year of the experience, P_{x+t} will consist of the lives under review a fraction of a year t after any 1st January who were of nearest age x on that 1st January. Again, if assumed age x is age x next birthday on any policy anniversary during the experience, P_{x+t} will be the number of lives who were under review a fraction of a year t after the anniversary on which they were aged x next birthday.

6. Distribution of deaths: alternative assumption.

It may be of interest to consider what would be the relation between m_x and q_x if, instead of assuming as in para. 2 an even distribution of the actual deaths, we assumed, as we did in earlier chapters when calculating E_x, that the deaths would be uniformly spread if there were no other movements.

To see the effect of this assumption, let us take a group of lives exposed from age $x+s$ to age $x+s+k$, there being no causes of increment or decrement other than death between these ages. Let P_x be the number of lives aged x of which P_{x+s} and P_{x+s+k} would be the survivors if death were the only force operating and θ_x the number of deaths which would have occurred between ages x and $x+1$ under the same conditions. We shall assume that the θ_x deaths would be evenly distributed.

Then the deaths between ages $x+s$ and $x+s+k$ would number $k\theta_x$ and, by formula (5), the central exposed to risk would become

$$\int_s^{s+k} P_{x+t}\,dt = \int_s^{s+k} (P_x - t\theta_x)\,dt = k\left(P_x - \frac{2s+k}{2}\theta_x\right).$$

The central rate of mortality for the group is therefore

$$\frac{\theta_x}{P_x - \frac{2s+k}{2}\theta_x}.$$

The corresponding expression if the lives are exposed for the whole year of age is $\dfrac{\theta_x}{P_x - \frac{1}{2}\theta_x}$ and the two expressions are not identical unless $2s+k=1$, i.e. unless the fractional period of exposure has $x+\frac{1}{2}$ as its middle point.

The lives contributing to an experience can be divided into groups such as the above, of which the essential feature is that within the age limits imposed there are no movements but deaths. A group would comprise, for example, the lives who enter at age $x+t$ and remain under observation till age $x+1$ or previous death, in which case s and k would have the values t and $1-t$ respectively. The principal group would, of course, consist of the lives exposed from age x to age $x+1$ or previous death, and for this group $s=0$ and $k=1$.

As the value of m_x is not the same for each group, it follows that the value of m_x for the whole experience depends on the distribution of the exposed to risk over the year of age. In this respect m_x differs from q_x, for, assuming an even distribution of deaths and no movements except deaths, the rate of mortality experienced by the group under review from age $x+s$ to age $x+s+k$ on the basis adopted in Chapter II would be $\dfrac{k\theta_x}{kP_x}=\dfrac{\theta_x}{P_x}$. The rate is therefore the same for all values of s and $s+k$ between 0 and 1, and the rate for the whole experience is not dependent on the distribution of the exposed to risk over the year of age.

In the above demonstration we have assumed that the data are homogeneous and the groups large enough to make statistical errors negligible.

Clearly the assumption made in para. 2 is more convenient in practice, although it is, of course, inconsistent to use a different assumption from that made in calculating E_x. This objection is largely academic, however, as the difference between the assumptions is not of any practical significance.

7. Calendar year method of calculating m_x.

In Chapters IV and VI we noted the advantages of the calendar year method where a continuous record of the ages of a group of lives is kept. We shall now show how formula (5) can be used to produce a simple variation of this method.

Consider an experience extending over a single calendar year. Let P_x^0 be the number of lives of assumed age x on 1st January and P_{x+t}^t the number of lives under review a fraction t of a year later who were of assumed age x on 1st January. P_{x+t}^t will consist of the survivors of P_x^0 increased by new entrants and decreased by deaths and withdrawals. If we assume that, for $0 \leqslant t \leqslant 1$, P_{x+t}^t is a first degree function of t, formula (5) becomes

$$E_x^c=\int_0^1 (P_x^0-tk)\,dt,$$

where k is constant. This integral reduces to $\frac{1}{2}(P_x^0+P_{x+1}^1)$, where

P^1_{x+1} is the number of lives of assumed age $x+1$ at the end of the year, and we have

$$m_x = \frac{\theta_x}{\frac{1}{2}(P^0_x + P^1_{x+1})}. \qquad \text{......(6)}$$

To apply this method we need only know the number of lives at the beginning and end of the year classified according to age and the number of deaths classified according to age at the beginning of the year. The method can be extended to cover several calendar years, in which case the number of lives at the beginning of each year classified according to age will be required. Thus, for an investigation covering three years, the formula is

$$m_x = \frac{\theta_x}{\frac{1}{2}(P^0_x + P^1_{x+1} + P^1_x + P^2_{x+1} + P^2_x + P^3_{x+1})} \qquad \text{......(7)}$$

where P^n_x is the number of lives of assumed age x, n years from the beginning of the period, and θ_x is the number of lives dying during the period who were of assumed age x on the 1st January before death.

The assumption that P^t_{x+t} is a first degree function implies that the net increase or decrease in the number of lives under review (i.e. the difference between the number of new entrants on the one hand and deaths and withdrawals on the other) is evenly spread over the calendar year. If there are no new entrants or withdrawals, the assumption implies an even distribution of deaths. In general, the assumption is sufficiently close to the truth for practical purposes, but cases do arise where a significant error is involved.

Formula (6) leads to

$$q_x = \frac{\theta_x}{\frac{1}{2}(P^0_x + P^1_{x+1} + \theta_x)}.$$

Exactly the same result would be obtained by using an ordinary calendar year method and assuming an even distribution of new entrants and withdrawals, for, if n_x and w_x are the number of new entrants and withdrawals who were of assumed age x on the 1st January before entry and exit respectively, then

$$E_x = P^0_x + \tfrac{1}{2}n_x - \tfrac{1}{2}w$$

and, since $$P^1_{x+1} = P^0_x + n_x - w_x - \theta_x,$$

$$E_x = \tfrac{1}{2}P^0_x + \tfrac{1}{2}(P^0_x + n_x - w_x) = \tfrac{1}{2}(P^0_x + P^1_{x+1} + \theta_x).$$

The method of this paragraph is not therefore new, but is merely a new way of demonstrating the calendar year method when certain specific assumptions are made regarding the distribution of the movements.

We refer hereafter to the method of this paragraph as the *modified calendar year method.*

CHAPTER X

THE CENSUS METHOD

1. Theory of the census method.

In obtaining formulae (4) and (5) of the previous chapter we regarded E_x^c as made up of an infinitely large number of infinitesimal periods of exposure applicable to different ages between x and $x+1$. For example, in the case where P_{x+t} is the number of lives attaining exact age $x+t$ during the period of the investigation, $P_{x+t}\Delta t$ represents the total length of the periods of exposure falling between exact ages $x+t$ and $x+t+\Delta t$ at all points of time within the limits of the investigation. The integration gives the sum of these totals for all ages between x and $x+1$. Clearly we shall obtain the same result if we reverse the processes of summation, i.e. if we sum the periods of exposure between ages x and $x+1$ separately for each point of time and take the total of these summations.

If, therefore, the experience covers n years and ${}^sP_x$ is the number of lives aged between x and $x+1$ at a point of time s years from the beginning of the investigation period, we have

$$E_x^c = \int_0^n {}^sP_x\,ds. \qquad \ldots\ldots(1)$$

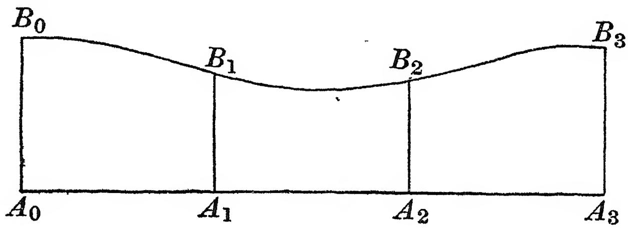
The central rate of mortality at exact age x is then given by the expression $\frac{\theta_x}{E_x^c}$, where θ_x is the number of deaths between exact ages x and $x+1$ during the period of the investigation.

Since ${}^sP_x$ is the number of lives who would have been recorded at age x last birthday if a census had been taken at time s, this method of obtaining the central death rate is generally known as the *census method* and is, in fact, the method normally used to obtain mortality rates from national statistics such as census and death returns.

2. Practical application of the method.

At first sight, this alternative method of obtaining E_x^c does not appear helpful, for ${}^sP_x$ cannot normally be expressed as a function of s which lends itself to integration. By having recourse to approximate methods, however, results of very great practical importance can be obtained.

Consider the following diagram, which represents an experience covering three years.



The time since the beginning of the investigation is measured along the line A_0A_3 from A_0; the ordinates represent the values of ${}^sP_x$ as s varies from 0 to 3, so that $B_0B_1B_2B_3$ is the graph of this function. The particular values shown are those applicable to integral values of s, i.e. the number of lives aged between exact ages x and $x+1$ at the beginning of the period and at the end of each complete year.

E_x^c is represented by the area $A_0A_3B_3B_0$ and, by calculating a sufficient number of values of A_sB_s, this area can be estimated with considerable accuracy. The number of values required will, of course, depend on the extent and nature of the data, i.e. on the form of the curve $B_0B_1B_2B_3$.

Dealing for the moment with a single year, we may assume, in the absence of any information to the contrary, that over this comparatively short period the population at any age increased or decreased uniformly, i.e. that the graph of ${}^sP_x$ is a straight line. The area $A_0A_1B_1B_0$ will then become a quadrilateral of unit base and its area will be the mean of the initial and final ordinates; hence

$$E_x^c = \int_0^1 {}^sP_x\,ds = \tfrac{1}{2}({}^0P_x + {}^1P_x) = {}^{\frac{1}{2}}P_x. \qquad \ldots\ldots(2)$$

For a period of three years

$$E_x^c = \tfrac{1}{2}{}^0P_x + {}^1P_x + {}^2P_x + \tfrac{1}{2}{}^3P_x$$

and for an investigation covering n years

$$E_x^c = \tfrac{1}{2}{}^0P_x + \sum_{s=1}^{n-1} {}^sP_x + \tfrac{1}{2}{}^nP_x. \qquad \ldots\ldots(3)$$

Formula (3) expresses E_x^c in terms of annual values of ${}^sP_x$, i.e. in terms of the results of annual censuses of the exposed population. This is convenient when we are dealing with life office data, since the information is immediately available from the office records. (The valuation statistics, for example, if lives are grouped according to year of birth, give us what we require at once.) For national tables, the problem is not so simple, since censuses are usually taken only at intervals of several years—in Great Britain the normal period is ten years. In these circumstances, there are two courses open to us. If the period of the investigation is short, say three years or less, we may base the value of the integral

$$\int_0^n {}^sP_x ds$$

on a single value of ${}^sP_x$. If, on the other hand, the investigation covers the whole of an intercensal period of ten years, the exposed to risk will be based on the populations at the two censuses. When so long a period is involved, the assumption that ${}^sP_x$ can be represented by a straight line is no longer tenable and various methods have been adopted to obtain a more accurate representation. These and other practical problems are discussed in Chapter XX, which deals with the construction of mortality tables from national statistics.

The census method as outlined above is not restricted to populations grouped according to age last birthday. It is applicable in just the same way if ${}^sP_x$ represents the number of lives of *nearest* age x at point of time s and θ_x the number of lives dying at nearest age x, in which case the resulting rate will apply to age $x-\frac{1}{2}$. The one essential is that the age grouping should be the same for both populations and deaths. A good rule to keep in mind is that a death at point of time s should be included in θ_x only if the life in question would have been included in ${}^sP_x$ if still alive.

3. Comparison with the modified calendar year method.

The differences between the census method and the method of (IX, 7) are important and should be studied with great care. In the one case E_x^c corresponds to the deaths between ages x and $x+1$; in the other to the deaths aged x at the beginning of the calendar year of death. The census method is thus a life year, the other a calendar year method.

In both cases we have assumed that a particular function increases or decreases in arithmetic progression. In the modified calendar year method the function is based on the same body of lives throughout the calendar year, increased by new entrants and reduced by deaths and withdrawals. On the other hand, in the census method the function ${}^sP_x$ represents a body of lives of which the composition is constantly changing throughout the calendar year as new lives are brought in by reaching exact age x and others withdraw by attaining exact age $x+1$. The progression of ${}^sP_x$ is therefore dependent on the age distribution of the lives contributing to the data between exact ages x and $x+1$. This feature of the census method introduces certain considerations which we must discuss further.

4. Investigation of underlying assumptions.

${}^sP_x$ may be divided into two groups of lives—those who were aged between x and $x+1$ at the beginning of the calendar year and those who, if they survive, will be between those ages at the end of that year. At the beginning of the year, all the lives in ${}^sP_x$ fall into the first group; this group however reduces throughout the year until at the end of the year all the lives in ${}^sP_x$ are in the second group. If we confine our discussion to the first year of the investigation—as we can do without loss of generality—we can therefore say that the rate at which the first group diminishes as s increases depends mainly on the age distribution of ${}^0P_x$ and, similarly, that the rate of increase in the second group depends mainly on the age distribution of ${}^1P_x$.

The distribution of the new entrants, deaths and withdrawals also has an effect on the rates of decrease and increase, but this factor can generally be neglected unless the number of such

movements is large and their distribution over the calendar year unequal. Neglecting this factor, we can say that the census method is strictly correct provided that either

(*a*) the ages of the lives in both ${}^0P_x$ and ${}^1P_x$ are evenly spread over the year of age x to $x+1$, in which event

$${}^sP_x=(1-s)\,{}^0P_x+s\,{}^1P_x,$$

or (*b*) $${}^0P_x={}^1P_x$$

and the distribution of the ages of the lives, though not necessarily uniform, is the same in both cases; ${}^sP_x$ will then be constant, for each life passing out of the exposed to risk by attaining age $x+1$ will at once be replaced by a life attaining age x.

Clearly, the assumption of para. (2) that ${}^sP_x$ can be represented by a straight line is accurately fulfilled if either proviso is satisfied. In all other cases, this assumption involves an approximation whose extent we can estimate by comparing the magnitudes and age distributions of ${}^0P_x$ and ${}^1P_x$. In this way we can decide whether or not the data are suitable for the application of the census method.

As an example, let us consider a group of lives aged x last birthday on 31st December $1900+n$, sufficiently numerous for their age distribution to be free from large random errors. Then, assuming the data to apply to Great Britain, we should normally expect the distribution of birthdays in the group to conform fairly closely to the distribution of births in Great Britain in the year $1900+n-x$. We can therefore test the suitability of the census method to obtain E_x^c in the year $1900+n$ by comparing the births of the years $1899+n-x$ and $1900+n-x$. In general, the number of births in Great Britain in the middle two quarters is greater, but only slightly greater, than the number in the first and last quarters. The general trend is the same in most years and the numbers in adjacent years do not usually differ much. In most years, therefore, conditions (*a*) and (*b*) are both satisfied sufficiently closely to make the error involved in the census method insignificant for practical purposes.

Exceptions, of course, occur. The war of 1914–18, for example, caused wide fluctuations in the birth-rate. At the end of the war

the rate rose rapidly, reached a peak in the first quarter of 1920 and decreased rapidly thereafter. In consequence, the age distributions of the survivors of lives born in 1919 and 1920 respectively are dissimilar and data drawn from these generations are not usually suitable for the census method. The error involved in applying the method in the year $1920+x$ to lives aged x last birthday is approximately 6 per cent. We shall see in a later chapter how the difficulty has been overcome in constructing national tables.

5. Application of the census method to life office data.

In a life office investigation we may expect the distribution of birthdays of the lives born in a particular year to conform to the distribution for the same generation of the general population. The arguments of the previous paragraph will therefore apply.

The shape of the distribution will, however, be liable to distortion by random errors due to paucity of data. An error of this kind is less serious than an error inherent in the formula, for the latter will usually operate in the same direction at adjacent ages, whereas the former will not operate according to a defined plan, Random errors from this source are, in fact, akin to errors in the numerator of m_x due to random variations in the number of deaths. The process of graduation is intended to eliminate random errors and the fact that there is an additional source of such errors in the census method increases the difficulty of achieving a successful graduation. As however the errors due to random variations in the distribution of births will, on the average, be small compared with those due to random variations in the number of deaths, the difficulties will not be seriously aggravated.

We conclude, therefore, that the census method applied to life office data should give results nearly as accurate as those obtained by earlier methods, provided that the distribution of births of the general population in the appropriate years was, roughly speaking, normal.

It may be asked what is the advantage of the census method for life office data, compared with the modified calendar year method of Chapter IX, which does not involve any assumption as to the

distribution of births. In the case of national statistics the advantages are readily apparent, for the official records of deaths give only the age at death; deaths cannot therefore be grouped according to their ages on the previous 1st January. This argument does not, however, apply to a life office experience, where full information is available.

For aggregate data, the modified calendar year method is preferable. Both methods are simple to apply inasmuch as they do not involve the preparation of individual cards for the contributing lives or the tabulation of movements other than deaths. The only advantage of the census method lies in the fact that it is a little more convenient to classify deaths by age last birthday rather than according to age on the preceding 1st January. This advantage is, however, more than outweighed by the greater accuracy of the modified calendar year method.

The justification for the census method lies in the fact that, as we shall see in Chapter XII, it can be used to calculate select rates, while the calendar year method is unsuitable for this purpose. In practice, select rates are almost invariably required in a life office investigation.

6. Alternative demonstration of the census method.

If the alternative form of formula (2) is adopted, i.e. $E^c_x = {}^{\frac{1}{2}}P_x$, we have $m_x = \dfrac{\theta_x}{{}^{\frac{1}{2}}P_x}$ and $q_x = \dfrac{\theta_x}{{}^{\frac{1}{2}}P_x + \frac{1}{2}\theta_x}$. An alternative demonstration of the census method which is sometimes given leads direct to this formula for q_x. This demonstration proceeds by the following steps:

(*a*) If we have a group of lives aged $x+\frac{1}{2}$ exactly on 1st July in a particular year ($P'_{x+\frac{1}{2}}$), life years will coincide with calendar years and all the corresponding deaths between ages x and $x+1$ will fall in the calendar year in question.

(*b*) Assuming that there are no movements except deaths and that the deaths are evenly spread over the year, the number aged x exactly on 1st January must have been $P'_{x+\frac{1}{2}} + \frac{1}{2}\theta_x$, whence

$$q_x = \frac{\theta_x}{P'_{x+\frac{1}{2}} + \frac{1}{2}\theta_x}.$$

(*c*) These conditions cannot arise in practice, but if we assume that the population between ages x and $x+1$ on any date in the calendar year is aged $x+\frac{1}{2}$ on the average and that ${}^sP_x$ takes the form of a straight line $(0 \leqslant s \leqslant 1)$, then ${}^{\frac{1}{2}}P_x$ is the mean population and we can take the average age to be $x+\frac{1}{2}$.

(*d*) On these assumptions, ${}^{\frac{1}{2}}P_x$ may be treated as if it were the number of lives aged $x+\frac{1}{2}$ exactly on 1st July and the argument in (*a*) and (*b*) above may be used to obtain the relation

$$q_x = \frac{\theta_x}{{}^{\frac{1}{2}}P_x + \frac{1}{2}\theta_x}.$$

This process is sound enough, but it has several disadvantages as a demonstration of the census method. In the first place, the underlying assumptions are not so readily apparent as in the earlier demonstration. Secondly, the reasoning behind the process may easily be misunderstood.

To avoid misunderstanding, it is desirable that, once we have decided to treat ${}^{\frac{1}{2}}P_x$ as a group of lives aged $x+\frac{1}{2}$ exactly and θ_x as the corresponding deaths between ages x and $x+1$, we should banish from our minds any thought of calendar years and concentrate on years of life. Failing this, there is a danger of our getting the impression that a direct connection exists between ${}^{\frac{1}{2}}P_x + \frac{1}{2}\theta_x$ and ${}^0P_x$, although in fact the former is an approximation to the number of lives aged x exactly which would give rise to θ_x deaths between exact ages x and $x+1$ and the latter is the number of lives aged between x and $x+1$ on a particular date. In other words, the latest demonstration may suggest erroneously that we are tracing a group of lives throughout the calendar year as in the calendar year method, whereas our data really consist of populations between ages x and $x+1$, the members of which at the beginning and end of the calendar year are entirely different.

7. Second alternative demonstration.

There is yet another way of demonstrating the census method which may be of interest. Assuming that the birthdays of the lives included in ${}^0P_x$ and ${}^1P_x$ are evenly distributed over the calendar year, the average contribution to E_x^c is $\frac{1}{2}$, if we exclude new

entrants, deaths and withdrawals. Hence we obtain formula (2), subject to any adjustment necessary on account of movements between ages x and $x+1$.

The assumption of an average exposure of $\frac{1}{2}$ is an overstatement in the case of those lives included in ${}^0P_x$ who passed out of observation by death or withdrawal before attaining age $x+1$. On the other hand, formula (2) makes no allowance for lives who attained age x during the year but passed out of observation before the end of the year.

For new entrants at age x last birthday during the year, the average exposure is less than $\frac{1}{2}$ for those who were still aged x last birthday at the end of the year. On the other hand, formula (2) does not include any exposure in respect of those who attained age $x+1$ before the end of the year.

If we assume that the net increase or decrease due to new entrants, deaths and withdrawals takes place uniformly over the calendar year, there will be a rough balance between the overstatements and understatements caused by our including an average exposure of $\frac{1}{2}$ for those under review at the beginning or end of the year and no exposure at all for the others. In general, the assumption is sufficiently correct for practical purposes and no adjustment of formula (2) is therefore necessary when the experience is extended to include movements at age x last birthday.

This proof is not so clear as the original proof, since the nature of the underlying assumptions is less readily apparent, but it serves as an introduction to the argument of the next paragraph.

8. Experience covering more than one calendar year.

The liability to error due to the age distribution may be considerably smaller when the experience covers several years.

From our definition of E_x^c, we see that the term $\sum_{s=1}^{n-1} {}^sP_x$ in formula (3) gives the correct exposure for all the lives contributing to the experience between their birthdays in the first and last years, except those who entered, died or withdrew between ages x and $x+1$. Moreover, the formula is approximately correct for

these lives also, if we assume, as we did in the last paragraph, that there is a rough balance of overstatements and understatements.

There remains the exposure before the birthday in the first year of the experience and after the birthday in the last year against which we have the term $\frac{1}{2}({}^0P_x + {}^nP_x)$ in formula (3). Although these data relate to two calendar years, they can be regarded in the same way as if ${}^0P_x$ and ${}^nP_x$ were the numbers aged x last birthday at the beginning and end of the same calendar year. If we denote by ${}^s\overline{P}_x$ the total number from the two groups aged x last birthday at time s from the beginning of the appropriate year, so that ${}^0\overline{P}_x = {}^0P_x$ and ${}^1\overline{P}_x = {}^nP_x$, a sufficient condition for the correctness of the term $\frac{1}{2}({}^0P_x + {}^nP_x)$ is that ${}^s\overline{P}_x$ should be a first degree function of s $(0 \leqslant s \leqslant 1)$.

We see, therefore, that the correctness of formula (3) depends primarily on the size and age distribution of the populations at the beginning and end of the experience. Suppose that the investigation relates to the years 1925–34 and that the distribution of births in earlier years has followed the normal pattern except in the years 1916–21. Let us consider, for example, the exposed to risk at age 10. ${}^0P_{10}$ and ${}^{10}P_{10}$ relate to the years of birth 1914 and 1924 in both of which years the birth distribution was, for practical purposes, uniform. Hence ${}^s\overline{P}_{10}$ would take the form of a straight line and formula (3) could be used to calculate E_{10}^c, in spite of the fact that lives born in the years 1916–21 attained age 10 during the investigation period.

At age 15 the position is different. ${}^0P_{15}$ and ${}^{10}P_{15}$ relate to the years of birth 1909 and 1919 and the distribution of births in the latter year was abnormal. We should therefore have to investigate the matter more fully before applying formula (3).

It should be noted that formula (3) is applicable only if we know the population at the end of each year. The argument of this paragraph is not therefore of interest when we are considering national statistics obtained from decennial censuses.

9. Relation between deaths and exposed to risk.

One feature which the reader may find difficult to understand is the fact that a life may contribute to the numerator of m_x, but

not to the denominator. If, for example, the experience related to the year 1940 and a life attained age 25 on 1st July 1940 and died on 1st October 1940, a unit would be included in θ_{25}, but nothing in either ${}^0P_{25}$ or ${}^1P_{25}$. At first sight, this seems to be contrary to the principle laid down in Chapter II that we must not include in the numerator deaths for which no exposure has been included in the denominator. It must be remembered, however, that the denominator $\frac{1}{2}({}^0P_{25}+{}^1P_{25})$ is an approximation to the integral $\int_0^1 {}^sP_{25}\,ds$. If the integral were evaluated exactly, the life in question would contribute $\frac{1}{4}$ to E^c_{25}, since it would be included in ${}^sP_{25}$ for all values of s between $\frac{1}{2}$ and $\frac{3}{4}$. It is quite unsound to test the correctness of the approximate expression substituted for the integral by examining the contribution made by an individual life.

10. Distribution of deaths over the year of age.

The assumptions on which the census method is based relate to the progression of the function ${}^sP_x$ over the calendar year and it is not necessary to make any assumption about the distribution of the deaths over the year of age. The method does not therefore introduce an error in m_x of the type referred to in (IX, 3). If, however, q_x is obtained from m_x by the relation $q_x=\frac{2m_x}{2+m_x}$ or by the formula $q_x=\frac{\theta_x}{E^c_x+\frac{1}{2}\theta_x}$, an error of this type will result. At old ages, where the error may be significant, the deaths between ages x and $x+1$ according to the mortality table will on the average occur before $x+\frac{1}{2}$; E_x will therefore be understated and q_x overstated by taking $\frac{1}{2}\theta_x$ as the difference between E_x and E^c_x or by using the normal relation connecting q_x and m_x. The error will be of the same magnitude as the error in m_x in (IX, 3), but its sign will be reversed.

An approximate adjustment could be made by adding to E^c_x a unit for each death between ages x and $x+\frac{1}{2}$ instead of half a unit for each death between ages x and $x+1$. In practice, however, it

is customary to neglect the error in an experience of assured lives, seeing that it causes an overstatement of q_x and is therefore on the side of safety when the table is used for calculating premiums. In an experience of annuitants, on the other hand, the error is not on the safe side and, moreover, the rates of mortality at advanced ages are far more important than in the case of an experience of assured lives. The effect of the error should therefore be examined in order to see whether it is sufficiently important to require adjustment of the rates at the old ages.

CHAPTER XI

CONSTRUCTION OF SELECT TABLES

1. Hitherto we have assumed that the rate of mortality depends only on the attained ages of the lives exposed to risk; in the methods and formulae we have so far devised all lives of the same age are accordingly grouped together. Mortality tables constructed in this form are known as *aggregate tables*.

It is well known, however, that in many of the mortality experiences with which the actuary is concerned, the rate of mortality is a function of variables other than the attained age. In an experience of assured lives, for example, the mortality will almost certainly vary not only with the age attained but with the period which has elapsed since the policy was effected and the life assured accepted as a first-class life. As long ago as 1834, Arthur Morgan recognized this and wrote in his preface to the mortality experience of the Equitable Society for the years 1762–1829:

"In a body of lives of the same age, all selected as healthy from the general mass of mankind, it is obvious that the rate of mortality must be considerably less for the first ten or twenty years after selection, than amongst those from whom they are chosen. As, however, these selected lives advance in age, their general health and the rate of mortality amongst them will naturally approximate to the common standard....The correct method, therefore (if sufficient data existed), would unquestionably be to make distinct tables from the mortality of each distinct class."

Tables of the type envisaged by Morgan have since become familiar to actuaries under the name of *select tables*; the feature of the experience which they reflect is similarly referred to by the name *selection* which he gave to it. Actually, other forms of selection exist and the present type is more accurately described as *temporary initial selection*. We must now consider the nature of this selection and the methods of constructing the select tables for which it calls.

2. Temporary initial selection.

This may arise from any particular incident or circumstance which produces a temporary variation in mortality. From the point of view of life assurance it is of the first importance, for, as indicated in the previous paragraph, it occurs to a marked degree when an office, in considering proposals for life assurance, declines or accepts on special terms those proposers whose prospects of longevity are below a certain standard. The remaining proposers who are accepted at ordinary rates thereby become a *selected* class, whose mortality may be expected to show temporary initial selection and who are generally referred to as *select lives*. Let us consider the effect of those lives on the mortality experienced by the office.

The lives contributing to the experience at age y will have been selected at different ages in the past. If x represents the age at entry and t the period which has elapsed since entry, usually called the duration in force, then $y=x+t$, where t ranges from 0 upwards and x ranges from y downwards. Now consider two groups of lives aged y, for the first of which $t=0$ and $x=y$ and for the second $t=10$ and $x=y-10$. If we examined separately the mortality of the two groups between ages y and $y+1$, we should expect the first group to show lower mortality than the second, since all the lives included in it were apparently in good health at age y when their proposals were submitted. The lives in the second group were similarly classed as healthy at age $y-10$, but it is unlikely that all of them would have been so classed if they had submitted new proposals at age y. It is therefore probable that they will experience heavier mortality between ages y and $y+1$ than the lives in the first group.

In fact, if we consider the series of values of q_y corresponding to different durations t and entry ages x, we should expect, other things being equal, a gradual increase as t increases and x decreases. We should further expect the rate of increase of q_y to slow down as t increases until eventually a value would be reached above which duration would have no apparent effect on mortality. This is an example of temporary initial selection and, as a result of its operation, the data for age y, were we to deal with all lives of the same attained age together, would be heterogeneous. To attain homogeneity, we

should have to subdivide the data according to duration in force as well as attained age, calculating separate rates of mortality for each subdivision.

3. Select tables: duration of selection.

Readers will be familiar with the usual form of select tables and with the use of the suffix $[x]+t$ to distinguish symbols relating to lives who entered t years ago at age x.

The effect of temporary initial selection, as the name implies, decreases with duration, and select tables are required only for those durations at which the effect is marked. The data applicable to longer durations are aggregated and mortality rates calculated therefrom regardless of duration. These are called *ultimate rates.*

The following table is taken from the $O^{[M]}$ experience and shows the values of $q_{[40-t]+t}$ for values of t between 0 and 10:

Duration t	Age at entry $40-t$	$q_{[40-t]+t}$
0	40	·00438
1	39	·00637
2	38	·00719
3	37	·00770
4	36	·00812
5	35	·00849
6	34	·00883
7	33	·00915
8	32	·00945
9	31	·00973
10 and over	30 and under	·00986

This example shows a characteristic feature of most select tables, i.e. the rapid increase in the values of $q_{[y-t]+t}$ where t is small, and the gradual slowing down of the rate of increase as t increases. The rates in the example are not the crude rates obtained direct from the data, but the graduated rates. It is not to be expected that ungraduated rates will progress so smoothly.

It is not usually easy to distinguish when the effect of selection ceases, for the trend of the ungraduated values of $q_{[y-t]+t}$ may be obscured, except for the smaller values of t, by random errors due to paucity of data and by other factors which will be discussed in

a later chapter. In the past, select tables have been constructed with select periods varying from one to ten years, and in most cases there has been a good deal of controversy about the correctness of the period chosen.

To investigate how long the effect of selection lasts, the actuary will construct tables of ungraduated rates in the form of the foregoing table. The average rates for quinquennial age groups might be used instead of those for individual ages so as to reduce the effect of random errors, for it must be remembered that even in a large experience the exposed to risk at an individual age and duration may be small. The actuary must then decide upon the length of the select period to adopt for the mortality table. If the length of this period is underestimated, the data outside its range will be heterogeneous and the corresponding rates will have the same disadvantages as aggregate rates, although to a lesser degree. As against this, there are practical objections to a long select period, for the longer this is the greater will be the number of functions to be calculated before the table can be put to practical use. It is the task of the actuary to decide how much weight to give to each of these opposing factors. Speaking generally, considerations of practical convenience cannot justify the reduction of the select period to such an extent as to make the remaining data markedly heterogeneous. If this were done, difficulty would be experienced in effecting a smooth junction between the select and ultimate rates of mortality.

If, for example, the select period for the $O^{[M]}$ table had been limited to three years instead of ten, the ultimate rate, which would have been the weighted mean of the last eight values in the table on p. 101, might have been about ·00900 and the progression of $q_{[40-t]+t}$ would have been as follows:

Duration t	Age at entry $40-t$	$q_{[40-t]+t}$
0	40	·00438
1	39	·00637
2	38	·00719
3 and over	37 and under	·00900

The last value is clearly out of keeping with the trend of the other values, and the result of such discontinuities would be to cause inconsistencies in some of the monetary functions calculated from the table, quite apart from the actual errors caused by distortion of the rates of mortality.

It might be argued that, if the select period had been reduced by two or three years only, the gain from the standpoint of convenience in use would have justified the comparatively slight distortion which this small reduction would have caused. The reader need not concern himself with so controversial a topic at this stage, but it is just as well that he should realize that the length of the select period to be adopted does not settle itself automatically and that the final decision cannot always be justified on strictly orthodox lines.

4. Select rates of mortality: exposed to risk.

Before we enquire how the methods of Chapters IV and V can be applied to select data, it might be well at this point to formulate a definition of select rates of mortality.

The annual rate of mortality experienced at duration t by a group of lives who entered at age x is the ratio of the number of deaths occurring in the group between durations t and $t+1$ to the number of lives attaining duration t, there being no causes of increment or of decrement other than death between these durations.

This definition suggests that the best way to obtain the exposed to risk for select rates of mortality is the method of age grouping by event or policy-year method. In practice this proves to be so and, apart from the census method, the policy-year method is the only one we need discuss. (In paragraph 10 we shall examine in more detail the reasons why an exact age method fails when applied to select rates.)

It will be remembered that, in Chapter V, lives were grouped according to their assumed ages at the beginning of the policy year of exposure, the assumed age on any policy anniversary being obtained by adding the duration in force to the assumed age at entry. For a select table, a similar procedure can be followed, but

the lives of assumed age y must be kept in separate groups according to their durations and each group dealt with by itself.

The simplest way of grouping the data in the required form is to group the lives first according to x, the assumed age at entry, and secondly according to t, the duration in force. Since x, y and t are connected by the relation $y=x+t$, we thus obtain in a simpler way the same groups as if we had analysed the data first according to y, the age attained, and secondly according to t.

The cards prepared for the individual lives will be similar to those used when constructing an aggregate table by the policy-year method, except that it is no longer necessary to record the ages at the dates of coming under and going out of observation, but merely the age at entry and the curtate durations at the above dates. If we take as an illustration an experience covering the years 1930–34 inclusive, the various symbols can be defined and the exposed to risk formulae developed exactly as in Chapter V. The close similarity of this paragraph and paragraph 3 of Chapter V should be noted:

$b_{[x]+t}$ = number of beginners who entered at nearest age x and whose curtate duration was t on 1st January 1930.

$e_{[x]+t}$ = number of enders who entered at nearest age x and whose curtate duration was t on 31st December 1934.

$\left.\begin{matrix}\theta_{[x]+t}\\ w_{[x]+t}\end{matrix}\right\}$ = number of deaths and withdrawals in 1930–34 who entered at nearest age x and whose curtate duration was t on the date of exit.

$n_{[x]}$ = number of new entrants at nearest age x in 1930–34.

An exposed to risk formula can then be built up in the same way as before, except that we shall now be connecting $E_{[x]+t}$ and $E_{[x]+t-1}$ instead of E_x and E_{x-1}. Let s be the average period between the date of entry and the succeeding 31st December. Then $b_{[x]+t}$ are exposed on the average for a fraction of a year $1-s$ at duration t and for a full year at duration $t+1$. Hence $(1-s)\,b_{[x]+t}$ must be added to $E_{[x]+t-1}$ in the formula for $E_{[x]+t}$ and a further $sb_{[x]+t}$ to $E_{[x]+t}$ in the formula for $E_{[x]+t+1}$. Similarly, $(1-s)\,e_{[x]+t}$ must be deducted from $E_{[x]+t-1}$ and a further $se_{[x]+t}$ from $E_{[x]+t}$, as the enders $e_{[x]+t}$ are exposed on the average for a fraction s of a year

at duration t. For the withdrawals, let h be the average period between the date of withdrawal and the preceding anniversary of entry, so that $(1-h)\,w_{[x]+t}$ must be deducted from $E_{[x]+t-1}$ and $hw_{[x]+t}$ from $E_{[x]+t}$. Proceeding on similar lines for the movements at duration $t-1$, we obtain the formula

$$E_{[x]+t}=E_{[x]+t-1}+sb_{[x]+t-1}+(1-s)\,b_{[x]+t}-se_{[x]+t-1}-(1-s)\,e_{[x]+t}$$
$$-hw_{[x]+t-1}-(1-h)\,w_{[x]+t}-\theta_{[x]+t-1}. \quad \ldots\ldots(1)$$

This formula applies when $t\geqslant 1$. When $t=0$, there are no terms relating to movements at the next lower duration, but as the new entrants at nearest age x are exposed for a full year at duration 0, the term $n_{[x]}$ must be included. Hence we have

$$E_{[x]}=(1-s)\,b_{[x]}-(1-s)\,e_{[x]}+n_{[x]}-(1-h)\,w_{[x]}. \quad \ldots\ldots(2)$$

Fractional exposures for beginners and enders can be avoided by excluding data from outside the period between the policy anniversaries in 1930 and 1934. This is often done in practice. The exposed to risk formulae are then simpler in form (cf. V, 2).

The formulae for the rates of mortality are

$$q_{[x]+t}=\frac{\theta_{[x]+t}}{E_{[x]+t}} \quad \text{and} \quad q_{[x]}=\frac{\theta_{[x]}}{E_{[x]}}.$$

These rates apply to nearest age x at entry, and a test should be made to find the exact age to which this corresponds.

As premium rates are based on age next birthday at entry, this age will always be recorded in the books of the office and it will usually be convenient to group the lives in this way in the mortality investigation. The procedure will be the same as when the lives are grouped by nearest age at entry, but the equivalent exact age at entry will of course be different.

5. Alternative method of dealing with fractional exposures.

If preferred, movements may be scheduled according to nearest duration at the date of the movement instead of curtate duration. All movements other than deaths, which must still be grouped according to curtate duration at death, will then be treated as if they occurred at exact durations. The exposed to risk formula will then be

$$E_{[x]+t}=E_{[x]+t-1}+b_{[x]+t}-e_{[x]+t}-w_{[x]+t}-\theta_{[x]+t-1}, \quad \ldots\ldots(3)$$

where $b_{[x]+t}$ = number of beginners who entered at nearest age x and whose duration in force was t on the anniversary of entry nearest to 1st January 1930, and $e_{[x]+t}$ and $w_{[x]+t}$ are defined on similar lines, but $\theta_{[x]+t}$ = number of deaths in 1930–34 who entered at nearest age x and whose duration in force was t on the anniversary of entry before the date of death. In the special case where $t=0$, the formula becomes

$$E_{[x]}=b_{[x]}-e_{[x]}+n_{[x]}-w_{[x]}. \qquad \text{......(4)}$$

6. Ultimate data.

As mentioned in paragraph 3 above, it will generally be difficult to decide what the length of the select period should be. We may therefore require to investigate the rates of mortality according to duration for a period considerably longer than the period of selection finally adopted. Once, however, the select period has been fixed, the data outside its limits need not be subdivided according to duration, since they are for practical purposes homogeneous with regard to duration. If in our investigations into the select period we have calculated $E_{[y-t]+t}$ and $\theta_{[y-t]+t}$ for all durations involved in the experience, we can obtain the ultimate exposed to risk at age y from the formula

$$E_y=E_{[y-n]+n}+E_{[y-n-1]+n+1}+\dots,$$

with a similar expression for θ_y, n being the length of the select period. Hence $q_y=\theta_y/E_y$ gives the ultimate rate of mortality at age y.

To obtain E_y and θ_y, it is not, however, necessary to calculate $E_{[y-t]+t}$ and $\theta_{[y-t]+t}$ for all values of t greater than $n-1$, as the methods of Chapter V can be applied to ultimate as well as aggregate data. The definitions of the various functions must of course be modified to exclude movements at durations less than n exact. There will be no term for new entrants in the exposed to risk formula, but an additional term is required to allow for the fact that the ultimate data are fed by transfers from the select data as the lives attain exact duration n. Thus, if F_y be the number of lives attaining exact duration n during the investigation whose assumed

age at entry was $y-n$, the formula for the ultimate exposed to risk is

$$E_y = E_{y-1} + F_y + \phi_y,$$

where ϕ_y consists of terms in b, e, w and θ.

The values of F_y may be obtained by inserting on the cards the assumed age on transfer from select to ultimate in those cases where the event occurred during the investigation and by sorting and totalling these cards. This process is not essential, however, for F_y can be obtained from the values of the select functions for duration $n-1$. The lives contributing to $E_{[y-n]+n-1}$ will be included in F_y, with the exception of those who passed out of observation at curtate duration $n-1$. We can therefore calculate F_y by suitably adjusting $E_{[y-n]+n-1}$ for movements at this duration.

The beginners, $b_{[y-n]+n-1}$, were exposed for a fraction $1-s$ at duration $n-1$, and $sb_{[y-n]+n-1}$ must be added in order that a whole unit will be included for each of them in F_y. Proceeding on similar lines for other movements, we obtain the relation

$$F_y = E_{[y-n]+n-1} + sb_{[y-n]+n-1} - se_{[y-n]+n-1} - hw_{[y-n]+n-1} - \theta_{[y-n]+n-1}. \quad \text{......(5)}$$

The complete formula corresponding to formula (2) of Chapter V is therefore

$$\begin{aligned} E_y = {} & E_{y-1} + sb_{y-1} + (1-s)\, b_y - se_{y-1} - (1-s)\, e_y - hw_{y-1} - (1-h)\, w_y \\ & - \theta_{y-1} + E_{[y-n]+n-1} + sb_{[y-n]+n-1} - se_{[y-n]+n-1} - hw_{[y-n]+n-1} \\ & - \theta_{[y-n]+n-1}. \end{aligned} \quad \text{......(6)}$$

Similarly, the formula corresponding to formula (1) of Chapter V is

$$\begin{aligned} E_y = {} & E_{y-1} + b_y - e_y - hw_{y-1} - (1-h)\, w_y - \theta_{y-1} + E_{[y-n]+n-1} \\ & - hw_{[y-n]+n-1} - \theta_{[y-n]+n-1}. \end{aligned} \quad \text{......(7)}$$

In this case, b_y would be the beginners on their policy anniversaries in 1930 such that y equals the assumed age at entry plus the duration on the policy anniversary, excluding those lives for whom this duration is less than n; and similarly for e_y. It should be noted that movements at exact duration n are included in the ultimate functions.

If nearest instead of curtate durations are used, the ultimate functions will include movements at nearest duration n, some of

which actually occurred within the select period; but there will of course be a compensating factor due to the treatment of movements between exact durations $n-1$ and $n-\frac{1}{2}$ as if they had occurred at exact duration $n-1$. In the case of the deaths, the grouping will be according to curtate duration as before and θ_y will include deaths at all durations not less than n exact. The formula corresponding to (6) will accordingly be

$$E_y = E_{y-1} + b_y - e_y - w_y - \theta_{y-1} + E_{[y-n]+n-1} - \theta_{[y-n]+n-1}. \quad \ldots(8)$$

7. Mechanical processes of sorting and grouping data.

Let us now consider the mechanical processes by which the data can be analysed and grouped. For each movement during the select period we require the assumed age at entry and the curtate or nearest duration at the date of the movement. For movements outside the select period we require the assumed age either on the last anniversary of entry before the date of the movement or on the nearest anniversary of entry. As this will be most easily obtained by recording the assumed age at entry and adding the curtate or the nearest duration, it is convenient to record the latter items on all the cards, whether or not the lives in question contribute to the select experience.

The cards should therefore be designed to show the following:

(1) date of birth,

(2) date of entry,

(3) nearest age at entry (or age next or last birthday),

(4) curtate (or nearest) duration on coming under observation,

(5) mode of coming under observation (beginner or new entrant),

(6) curtate (or nearest) duration at exit,

(7) mode of exit (ender, withdrawal or death),

(8) age on last anniversary of entry before coming under observation (or on nearest anniversary),

(9) age on last anniversary of entry before exit (or on nearest anniversary).

(1) need not be recorded if (3) is directly available from the books of the institution. (8) and (9) would be recorded only for movements for which the duration shown in (4) or (6) $\geqslant n$, where

n is the length of the maximum period during which we intend to trace the effect of duration on mortality.

The sorting of the cards might be carried out as follows, the results being tabulated at each stage:

(*a*) sort according to age at entry, $[x]$, those cards on which the duration according to item $(4) < n$;

(*b*) in each group, separate from the remainder those who entered during the period of the investigation ($n_{[x]}$);

(*c*) sort the remainder according to duration at the beginning of the investigation ($b_{[x]+t}$);

(*d*) amalgamate the cards for entry age $[x]$;

(*e*) sort according to mode of exit (e, w or θ) the cards for which the duration at exit $< n$;

(*f*) sort each group according to duration at exit ($e_{[x]+t}$, etc.);

(*g*) sort the cards excluded from (*a*) according to age on coming under observation as given in item (8) (b_y);

(*h*) amalgamate all the cards in (*g*) irrespective of age on coming under observation and add them to the cards excluded from (*e*);

(*i*) sort according to mode of exit (e, w or θ);

(*j*) sort each group according to age at exit as given in item (9) (e_y, w_y or θ_y).

A separate schedule would then be prepared for each entry age as follows:

Entry age x				No. of new entrants $n_{[x]}$		
Duration t	$b_{[x]+t}$	$e_{[x]+t}$	$w_{[x]+t}$	$\theta_{[x]+t}$	$E_{[x]+t}$	$q_{[x]+t}$
0						
1						
2						
...						
...						
...						
$n-1$						

The number of new entrants would be obtained by counting the cards in the groups mentioned in (*b*) and the figures for the second to fifth columns by counting the cards in the groups mentioned in

(*c*) and (*f*). $E_{[x]}$ and $E_{[x]+t}$ would then be calculated by the appropriate formulae according as curtate or nearest durations had been recorded. For the ultimate data, a single schedule in the following form would be required:

Age y	b_y	e_y	w_y	θ_y	F_y	E_y	q_y

where F_y is obtained by relation (5) when curtate durations have been used and $F_y = E_{[y-n]+n-1} - \theta_{[y-n]+n-1}$ when nearest durations have been used. The figures for the second to fifth columns would be obtained by counting the cards in the groups mentioned in (*g*) and (*j*) and those for the sixth column from the schedules applicable to the select period. E_y would then be calculated by the appropriate formula.

8. Financial effect of using select tables.

As an illustration of the importance of select tables, we may consider the financial effect of a life office using an aggregate table to calculate monetary functions. The aggregate data are based partly on select and partly on ultimate lives, and the aggregate mortality rates therefore lie between the select and the ultimate. At young ages, few of the policies will have been in force for long and the aggregate rates will differ very little from the select. At the old ages, however, there will be few new entrants and the aggregate rates will be only slightly less than the ultimate, although appreciably higher than the select rates. At the middle ages, the number of select and ultimate lives contributing to the data will both be comparatively large and there will be an appreciable difference both between the aggregate and select rates and between the ultimate and aggregate rates.

Let us consider the effect on premium rates for whole-life policies, bearing in mind that the premium at any particular age depends on the rates of mortality at all older ages. For a young entry age, the use of an aggregate table will result in a slight overstatement of the

amount required to provide assurance cover at the young ages, a considerable understatement of the amount required at the middle ages, and a slight understatement of that required at the old ages. On balance, the premium will be undervalued. For an old entry age, there will be a considerable overstatement of the cost of the cover during the select period and a slight understatement thereafter so that the net effect will be an overstatement of the premium. Over the whole range of entry ages there will, therefore, be a gradual change over from an understatement to an overstatement of the correct rate of premium.

The effect on reserve values is more complicated owing to the greater number of variables and no simple general conclusions can be reached. An examination of the figures for the particular table must always be made.

9. Other types of temporary initial selection.

Select mortality tables are not confined to experiences of assured lives. Annuitants exercise a certain amount of temporary initial selection, since a person in poor health is unlikely to purchase an annuity which will cease on death. The influence is not so strong as that exercised by a life office, but it is sufficiently strong to make the construction of select tables desirable.

Temporary initial selection may be reversed, i.e. new entrants at age x may experience heavier mortality at that age than lives now aged x who entered earlier. *Reversed selection* would be encountered if the mortality of pensioners who had retired owing to bad health were investigated. A group of pensioners who had just retired at age x would include some who would be unlikely to survive more than a few years and others who might no longer be fit for work but whose prospects of longevity would be little impaired. The death of most of the former lives within t years would leave a body of lives aged $x+t$ whose mortality at that age would be lighter than that of a group aged $x+t$ who had just retired. Neglect of this factor might lead to serious errors in valuing annuities payable to pensioners who had retired through ill-health.

10. Exact-age method.

If all lives entered on their birthdays, $\theta_{[x]+t}$ would be the number of lives aged $x+t$ last birthday dying between exact durations t and $t+1$. This suggests that, when life and policy years do not coincide, $\theta_{[x]+t}$ should be defined in this way and $E_{[x]+t}$ as the number of years of exposure falling within (*a*) the year of age $x+t$ to $x+t+1$ and (*b*) the year of duration t to $t+1$, with the usual proviso regarding the treatment of deaths in the exposed to risk. An example will show how these definitions would operate.

Consider a life born on 1st February 1900 who entered on 1st May 1925 at exact age $25\frac{1}{4}$, and suppose that the investigation relates to the years 1930–34. We cannot classify the life according to the age at entry, for this is not an exact age and we wish to use the exact age method. We see, however, that the life was aged 29 last birthday between 1st January 1930 and 31st January 1930, and that during this period the curtate duration was 4. Hence the period in question must be allocated to $E_{[x]+t}$, where $x+t=29$ and $t=4$, i.e. to $E_{[25]+4}$. Between 1st February 1930 and 30th April 1930, we have $x+t=30$ and $t=4$, so that this period must be allocated to $E_{[26]+4}$ and so on. The contributions made to the exposed to risk would therefore be as follows:

x	t	Fraction included in $E_{[x]+t}$
25	4	$\frac{1}{12}$
26	4	$\frac{1}{4}$
25	5	$\frac{3}{4}$
26	5	$\frac{1}{4}$

and so on until exit.

The laborious nature of the process is obvious and the method is never used directly. Nevertheless, as we shall see in the next chapter, we may sometimes wish to approximate to the value of $E_{[x]+t}$ according to the above definition.

If a policy-year method were used and lives grouped according to the age nearest birthday at entry, the life in the example would contribute one-third to $E_{[25]+4}$ and a unit to each of $E_{[25]+5}$, $E_{[25]+6}$, etc., until the policy year of exit.

Another possible method would be to trace lives through exact years of life and to approximate to the duration. We might, for example, take the assumed duration on a birthday to be the nearest integral duration. The duration on 1st February 1929 for the life referred to above would then be taken as 4, so that the assumed duration on 1st January 1930 would be $4\frac{11}{12}$ and the life would contribute $\frac{1}{12}$ to $E_{[25]+4}$, a unit to $E_{[25]+5}$, etc. In other words, the life would be treated as if he had entered at exact age 25 on 1st February 1925, which is the nearest birthday to the true date of entry. Deaths would, of course, be grouped according to the nearest duration on the birthday before death.

The latter method might be called an exact-age and approximate-duration method, as opposed to the method described at the beginning of this paragraph which might be termed an exact-age and exact-duration method. Similarly, the policy-year method might be called an approximate-age and exact-duration method. The approximate-duration method is less suitable for practical purposes than the policy-year method. Moreover, examination of any select table will show that at most of the durations for which it has been constructed $q_{[x]+t}-q_{[x]+t-1}$ is greater than $q_{[x]+t}-q_{[x-1]+t}$, showing that duration has a greater effect on mortality than age. It is, therefore, better to group the data by exact duration and approximate age rather than by exact age and approximate duration.

11. Rates of withdrawal.

It may be appropriate here to make a brief reference to rates of decrement other than mortality rates. In general, the methods used in this and in earlier chapters are equally suitable for the calculation of rates of decrement such as the rate of withdrawal. In calculating the exposed to risk of withdrawal we should have to include withdrawals to the same extent as if withdrawal had not taken place, but the deaths would be treated as exposed up to the assumed age at death only. In other respects, the exposed to risk of withdrawal would usually be the same as the exposed to risk of death.

In an experience of assured lives, duration will generally have a greater effect than age on rates of withdrawal. These rates are

seldom required and when an investigation is necessary, a rough idea of the trend of the rates will probably be sufficient. It may therefore be possible to neglect the effect of age and calculate rates according to duration only. Failing this, it should be sufficient if the data are divided into two or three age groups. At the short durations, the rates will normally be found to increase up to a maximum which will probably be reached when $t=2$, 3 or 4. This feature is, no doubt, the result of proposers effecting policies which are too large for their means or of their being persuaded to insure their lives against their inclination. A proportion of these lives will very soon decide to surrender, but once they have been eliminated the rate of withdrawal will fall, rapidly at first and then more slowly, to a comparatively low level. As the age of the policy increases, so does the deterrent to surrender and a gradual decrease in the rate is therefore likely to continue at the longer durations.

In calculating rates of withdrawal from an experience of assured lives, we are only concerned with exits due to lapse or surrender and it is therefore necessary to treat maturities of endowment assurances as a separate type of exit.

Example 1.

An office has been issuing policies of a particular type since 1st January 1924 and you have been asked to investigate the lapse experience. Show how you would obtain the lapse rates from the data in the following table, assuming that they depend on duration only.

Curtate duration	Number of policies in force on 31st December 1933	Exits before 31st December 1933	
		By death	By lapse
0	210	4	30
1	190	5	140
2	180	6	100
3	170	5	75
4	160	4	50
5	150	3	30
6	145	3	20
7	140	2	10
8	130	2	5
9	120	1	2

The data are given in a rather unusual form, but they can easily be re-arranged.

Let e_t be the number in force on 31st December 1933 at curtate duration t,

θ_t ,, ,, of deaths at curtate duration t,

ω_t ,, ,, of lapses ,, ,, ,, t.

There are no beginners, as all policies are included from the dates when they were effected. The number of new entrants must therefore be equal to the total number of exits; hence

$$n=\sum_{t=0}^{9}(e_t+\theta_t+\omega_t)=2092.$$

The exposed to risk formula, assuming that the deaths are evenly distributed over the policy year and that the policy anniversaries of the enders are evenly distributed over the calendar year, is

$$E_t^w=E_{t-1}^w-\tfrac{1}{2}(e_{t-1}+e_t)-\tfrac{1}{2}(\theta_{t-1}+\theta_t)-\omega_{t-1}$$

and the lapse rate at duration t is given by $q_t^w=\dfrac{\omega_t}{E_t^w}$.

At duration 0, we have $E_0^w=n-\frac{1}{2}e_0-\frac{1}{2}\theta_0$. The working schedule is therefore as follows:

Duration t	$\frac{1}{2}(e_{t-1}+e_t)$	$\frac{1}{2}(\theta_{t-1}+\theta_t)$	ω_{t-1}	E_t^w	q_t^w
0	105	2	0	1985	·0151
1	200	4·5	30	1750·5	·0800
2	185	5·5	140	1420	·0704
3	175	5·5	100	1139·5	·0658
4	165	4·5	75	895	·0559
5	155	3·5	50	686·5	·0437
6	147·5	3	30	506	·0395
7	142·5	2·5	20	341	·0293
8	135	2	10	194	·0258
9	125	1·5	5	62·5	·0320
10	60	·5	2	0	

Alternatively, the exposed to risk could be obtained direct from the table in the question. At duration 9, as there are no policies in force

for longer durations, the only lives contributing are those included in e_9, θ_9 and ω_9 and the enders and deaths are exposed for only half a year on the average.

Hence $$E_9^w = \tfrac{1}{2}(120) + \tfrac{1}{2}(1) + 2 = 62\cdot 5.$$

At duration 8, the above lives will each contribute a full year of exposure, while e_8 and θ_8 will each contribute half a unit and ω_8 a whole unit.

Hence $E_8^w = 120 + 1 + 2 + \tfrac{1}{2}(130 + 2) + 5 = 194$ and so on.

CHAPTER XII

USE OF CENTRAL RATES OF MORTALITY FOR SELECT TABLES. APPLICATION OF THE CENSUS METHOD

1. The theory underlying the use of central rates of mortality for select tables is similar to the theory already developed for aggregate rates, and it is unnecessary to set it out again in detail. The same is true of the application of the census method. In this chapter frequent references will be made to the corresponding formulae and arguments in Chapters IX and X, and the reader is advised to consult these and to keep the parallel closely in mind.

2. Definition of $m_{[x]+t}$.

The *central rate of mortality at duration t for a group of lives who entered at assumed age x* may be defined as the ratio of the number of deaths between durations t and $t+1$ to the number of years' exposure to the risk of death between these durations, there being no causes of increment or decrement other than death and each death being treated as exposed only up to the exact duration at death. The analogy with previous definitions is obvious.

3. Formulae for $E^c_{[x]+t}$.

Exposed to risk formulae can be devised in the same way as the corresponding aggregate formulae (IX, 2). In general, the formula for $E^c_{[x]+t}$ will differ from that for $E_{[x]+t}$ by the substitution of $\frac{1}{2}(\theta_{[x]+t-1}+\theta_{[x]+t})$ for $\theta_{[x]+t-1}$ in the denominator.

The alternative formula, corresponding to formula (5) of Chapter IX, is

$$E^c_{[x]+t}=\int_t^{t+1} P_{[x]+r}\,dr, \qquad \text{......(1)}$$

where $P_{[x]+r}$ is the number of new entrants at assumed age x who attained exact duration r during the investigation period.

4. Census method.

Formula (1) is based on the policy-year method of age grouping. The census method, on the other hand, is based on the exact-age or life-year method. As we have already seen (X, 1), it approximates in the case of aggregate rates to the central exposed to risk corresponding to the deaths between exact ages x and $x+1$. For select rates we require an approximation to the exposed to risk corresponding to the deaths between exact ages $x+t$ and $x+t+1$ at curtate duration t. This exposed to risk consists of the exposure falling between the same limits as the deaths, i.e. at age $x+t$ last birthday, curtate duration t, and can be regarded as the sum of an infinitely large number of infinitesimal periods of exposure.

Hence, by a similar argument to that used in (X, 1), if ${}^sP^t_{x+t}$ is the number of lives aged $x+t$ last birthday of curtate duration t at a point of time s years after the beginning of the investigation period,

$$E^c_{[x]+t} = \int_0^n {}^sP^t_{x+t}\,ds, \qquad \ldots\ldots(2)$$

n being the number of years over which the experience extends.

Then $m_{[x]+t} = \dfrac{\theta_{[x]+t}}{E^c_{[x]+t}}$, where the numerator is the number of deaths during the n years at age $x+t$ last birthday whose curtate duration was t at the date of death. This is the central rate of mortality for exact age at entry x and exact duration t.

5. Practical application of the census method.

On the assumption that ${}^sP^t_{x+t}$ is a linear function of s over each separate calendar year of experience, formula (2) reduces to

$$E^c_{[x]+t} = \tfrac{1}{2}\left({}^0P^t_{x+t} + {}^1P^t_{x+t}\right) = {}^{\frac{1}{2}}P^t_{x+t}, \qquad \ldots\ldots(3)$$

when the experience covers a single year and

$$E^c_{[x]+t} = \tfrac{1}{2}\,{}^0P^t_{x+t} + \sum_{s=1}^{n-1} {}^sP^t_{x+t} + \tfrac{1}{2}\,{}^nP^t_{x+t} \qquad \ldots\ldots(4)$$

when the experience covers a period of n years (cf. formulae (2) and (3) of Chapter X).

To apply the method, we therefore require periodic censuses classifying the lives according to age and duration.

We have already seen that the principal use of select tables is for life office work. The valuation books will usually provide a continuous record based on age, and it is quite a simple matter to devise a system whereby the number on the books at any time, classified according to duration as well as age, will be readily obtainable. It is, of course, only at the shorter durations that the additional classification is necessary. The application of the census method to the calculation of select rates for life office records does not, therefore, present any difficulties and has obvious merits on account of its simplicity.

As with aggregate rates, the age grouping can be according to nearest age instead of age last birthday. The deaths will then have to be grouped by nearest age and curtate duration at death, and $m_{[x]+t}$ will be the central rate at duration t for entrants at age $x-\frac{1}{2}$. The rule to be remembered is that a death at time s should not be included in $\theta_{[x]+t}$ unless the life in question would have been included in ${}^sP^t_{x+t}$ if still alive.

6. Investigation of underlying assumptions.

We shall confine our investigation to a single calendar year and neglect the influence of deaths and withdrawals throughout the year, just as we did when considering the aggregate experience (X, 4). As compared with the latter investigation, we have to take into account duration as well as age. ${}^sP^t_{x+t}$ will be increased when a life aged $x+t$ last birthday attains duration t at point of time s or when a life of curtate duration t attains age $x+t$; and similarly for lives passing out of ${}^sP^t_{x+t}$. The movements with which we are concerned may be summarized as follows:

Increases: (a) life passing from ${}^sP^t_{x+t-1}$ to ${}^sP^t_{x+t}$.
(b) life passing from ${}^sP^{t-1}_{x+t}$ to ${}^sP^t_{x+t}$.

Decreases: (a) life passing from ${}^sP^t_{x+t}$ to ${}^sP^t_{x+t+1}$.
(b) life passing from ${}^sP^t_{x+t}$ to ${}^sP^{t+1}_{x+t}$.

When dealing with the aggregate function ${}^sP_x$, we divided the lives into two groups according as they were aged $x-1$ or x last

birthday at the beginning of the year. The lives contributing to ${}^{s}P^{t}_{x+t}$ during the year may have fallen into any one of four groups at the beginning of the year, ${}^{0}P^{t-1}_{x+t-1}$, ${}^{0}P^{t}_{x+t-1}$, ${}^{0}P^{t-1}_{x+t}$ and ${}^{0}P^{t}_{x+t}$ although not all the lives in the second and third groups will contribute; e.g. a life in the second group might reach duration $t+1$ before age $x+t$ and would therefore pass from ${}^{s}P^{t}_{x+t-1}$ to ${}^{s}P^{t+1}_{x+t-1}$.

An analysis of the conditions which would justify our assumption that ${}^{s}P^{t}_{x+t}$ is a first degree function of s accordingly involves certain difficulties. If we were to deal with the problem exhaustively, we should have to measure the combined effect of age and duration as s increases. For our purpose, however, it will be sufficient if we lay down separate conditions regarding the distributions of birthdays and policy anniversaries respectively as if the two factors were operating independently. On this basis, we can argue by analogy with (X, 4) that the following conditions, applied to the lives aged $x+t-1$ and $x+t$ last birthday at the beginning of the year, having curtate durations $t-1$ or t at that date, would suffice to make ${}^{s}P^{t}_{x+t}$ a first degree function of s:

(*a*) (i) distribution of birthdays uniform over the calendar year for each of the groups aged $x+t-1$ and $x+t$ last birthday respectively when the lives are grouped according to age, or

(ii) equal numbers and the same distribution of birthdays (not necessarily uniform) in each of the two age groups;

(*b*) (i) distribution of policy anniversaries uniform over the calendar year for each of the groups of curtate duration $t-1$ and t respectively when the lives are grouped according to duration, or

(ii) equal numbers and the same distribution of policy anniversaries (not necessarily uniform) in each of the two groups based on duration.

We saw in Chapter X that, in general, conditions (*a*) (i) and (ii) are fulfilled sufficiently closely for practical purposes, provided that the data are adequate and the distribution of birthdays over the calendar year does not differ widely from the normal pattern. We shall now consider in what circumstances the conditions relating to duration will be fulfilled.

7. Effect of variations in the distribution of entry dates over the calendar year on the function ${}^sP^t_{x+t}$.

The number of lives attaining duration t in a given year depends primarily on the number of entrants t years earlier, provided that t is small. Similarly, the distribution of the policy anniversaries depends on the distribution of entry dates t years earlier. Our investigation accordingly reduces to a comparison of the number and distribution of new entrants from year to year.

The number of new entrants to a life office depends upon the new business pressure exerted by the office, the state of trade, the political outlook, the level of premium or annuity rates, the rate of interest obtainable on investments and many other factors. The distribution over the calendar year is determined by the varying influence of these factors from day to day. Under stable conditions, the distribution would conform fairly closely to a standard pattern. For lives assured, as opposed to annuitants, the number tends to rise to a peak at the end of the office year, and the distribution is accordingly less uniform than the distribution of births.

Any abnormal variations in the controlling factors will upset the standard distribution. Such variations may have an immediate effect on the flow of new entrants and may cause a considerable change in both the number and distribution in successive years. For example, the War Loan conversion in 1932 brought about an immediate demand for annuities in the middle of the calendar year, and the outbreak of war in 1939 caused a sudden fall in the volume of new assurance business (see Example 1, p. 126).

Prima facie, therefore, it appears that conditions (*b*) (i) and (ii) are less likely to be satisfied than the corresponding conditions relating to the number and distribution of births.

8. Experience covering more than one calendar year.

The error in $m_{[x]+t}$ due to changes in the number and distribution of new entrants is usually smaller when the experience extends over several years instead of a single year.

The terms $\sum_{s=1}^{n-1} {}^sP^t_{x+t}$ in formula (4) provide for the inclusion of a unit for each life aged $x+t$ last birthday having a curtate duration of t at the beginning of any year except the first. If we neglect the

influence of deaths and withdrawals, as we did in (X, 8), we can say that each of the above lives was in fact exposed for a full year at duration t. This exposure did not, however, occur entirely at age $x+t$ last birthday, but extended to the adjacent ages. In this respect formula (4) overstates the exposed to risk. On the other hand, the formula makes no allowance for the periods of exposure at age $x+t$ last birthday and curtate duration t, which did not include a census date. We must therefore consider whether we are justified in assuming that these errors will balance.

To give a reasoned answer to this question we should have to carry out a tedious and rather complicated investigation which would be out of place in a text-book concerned primarily with basic principles. The reader is therefore asked to accept the statement that the terms $\sum_{s=1}^{n-1} {}^sP^t_{x+t}$ should in general give a close approximation to the central exposed to risk between the policy anniversaries in the first and last years of the investigation for exact age at entry x and exact duration t.

The correctness of the expression depends to some extent on the number and age distribution of the entrants at age $x-1$ last birthday and x last birthday respectively, but only to a very limited degree on the number and distribution of new entrants in successive calendar years. So far as this part of the formula is concerned, we need not therefore trouble about fluctuations in the number of new entrants such as we discussed in paragraph 7.

If we accept the above statement, we are left with the first and last terms in formula (4), and we have to consider whether they provide a sufficiently close approximation to the exposure falling either before the policy anniversary in the first year or after the policy anniversary in the last year of the experience. We can put the matter to the test by considering the function ${}^s\overline{P}^t_{x+t}$, consisting of the total number of lives aged $x+t$ last birthday of curtate duration t at time s after the beginning of either year, including only those who had not reached their policy anniversaries in the first year and those who had passed their policy anniversaries in the last year. This expression will be a linear function of s provided that certain conditions, similar to those stated in paragraph 6,

are satisfied. We need not concern ourselves with the question of age distribution, beyond satisfying ourselves that the distribution of births in the general population conformed approximately to the normal pattern in the appropriate years (see X, 4). The effect of the distribution of entry dates requires more thorough investigation.

To obtain the conditions corresponding to (*b*) (i) and (ii) of paragraph 6, we must consider the following two groups of lives, i.e. those aged $x+t-1$ and $x+t$ last birthday of curtate duration t at the beginning of the first year of the experience and those aged $x+t$ and $x+t+1$ last birthday of curtate duration t at the end of the last year. The two sets of conditions we require are then as follows:

(*a*) distribution of policy anniversaries over the calendar year uniform in both groups;

(*b*) equal numbers and the same distribution of policy anniversaries (not necessarily uniform) in each group.

Either of these conditions is sufficient, assuming that the conditions as regards age distribution are fulfilled, to ensure that ${}^{s}\overline{P}^{t}_{x+t}$ is a linear function of s and hence that the first and last terms in formula (4) are correct.

It follows that the principal test of the suitability of the census method for select data covering a period of years consists of an examination of the number and the distribution of policy anniversaries of the lives under review at the beginning and end of the period. Example 2 (p. 128) may help the reader to understand this paragraph.

9. Conclusions as to the suitability of the census method for select data.

If we apply the criterion laid down in the preceding paragraph, we find in practice that the error introduced is usually small (cf. Example 2, where a rapid and marked change in the flow of new entrants is assumed). This suggests that, provided that the experience extends over several calendar years and the data are adequate, the census method will usually give sufficiently accurate select rates of mortality.

When any doubt exists, the number of lives of the appropriate durations at the beginning and end of the investigation period and the distribution of their policy anniversaries should be examined. The size of the error which would be introduced by using the census method can then be estimated. Provided that the numbers at the beginning and end do not differ greatly, it is usually safe to conclude without further investigation that the distribution of policy anniversaries is approximately the same in each case and that the necessary conditions are sufficiently nearly satisfied. The shorter the investigation period, the more attention must be paid to any differences which exist.

In formulating our conclusions at the beginning of this paragraph, a proviso was made about the adequacy of the data. This is, of course, an essential requirement, whatever method is employed. We have already seen that, when considering aggregate rates, random errors due to paucity of data tend to be rather greater for the census method than for the earlier methods (see X, 5). When dealing with select data, we shall have additional errors of this kind arising from random variations in the distribution of entry dates. The errors from this source, like those due to random variations in the age distribution, will be small on the average compared with those, common to both the policy-year method and the census method, which arise from random variations in the number of deaths. We conclude, therefore, that, if the select data are adequate for the policy-year method, we need not be discouraged from using the census method on account of paucity of data.

In a later chapter we shall discuss a number of practical points arising out of the use of the method.

10. Modification of the census method.

The introduction of errors due to changes in the distribution of entry dates can be largely avoided by excluding part of the data. If we include in the numerator only the deaths occurring between the policy anniversaries in the first and last years of the experience, the corresponding exposed to risk will be given by $\sum_{s=1}^{n-1} {}^sP^t_{x+t}$ and, as stated in paragraph 8, the correctness of these terms is for practical

purposes independent of fluctuations in the number of new entrants.

If this modified method were to be used, it would be convenient to record for the deaths in the first and last years not only the curtate duration at death but also the curtate duration at the beginning of the first or end of the last year. $\theta_{[x]+t}$ would include only those deaths for which these two durations were different.

In general, however, the gain in accuracy by using the modified method is so slight as to make this small additional amount of work unnecessary.

11. Error at duration 0.

The greater part of the exposed to risk falling after the policy anniversary in the last year of an experience relates to the earlier part of the policy year, and the reverse is true of the exposure before the policy anniversary in the first year. If there is a considerable difference between the amounts of exposure in the two periods there will not be an even distribution of the exposed to risk over the policy year. The resulting error will be insignificant unless the force of mortality is increasing rapidly over the policy year, a state of affairs which seldom exists except, perhaps, at duration 0. The error introduced at that duration may be significant.

The possible existence of an error of this nature has been advanced as an objection to the census method. Actually, a similar error is inherent in the policy-year method, when the data relating to the periods before the policy anniversary in the first year and after the policy anniversary in the last year are included. The exclusion of these data would eliminate the error, but the same result can be achieved for the census method by using the modification explained in paragraph 10.

It should be noted that the error which we have just discussed is quite distinct from those mentioned earlier in the chapter.

12. Modified policy-year method.

We saw in paragraph 8 that $\sum_{s=1}^{n-1} {}^sP^t_{x+t}$ may be regarded as an approximation to the central exposed to risk between the policy

anniversaries in the first and last years for exact age at entry x and exact duration t. Alternatively, the expression may be regarded as an approximation to the central exposed to risk at exact duration t for entrants at assumed age x, where assumed age x means age x last birthday on 1st January following entry. The correct exposed to risk for these lives would differ from the above expression only in respect of exits at curtate duration t, and the error involved in the approximation would usually be negligible. This is, of course, a policy-year method and the deaths corresponding to the above exposed to risk are those occurring at curtate duration t among lives entering at assumed age x. The age at death will not always be $x+t$ last birthday as in the census method.

The resulting rates are not necessarily those for exact entry age x, and a sample of the data might have to be examined to ascertain the exact age at entry corresponding to assumed age x.

Some actuaries have indicated that they prefer the modified policy-year method to the census method on account of its greater accuracy. This view has been answered by others, who consider that the small gain in accuracy is outweighed by the additional work of scheduling the deaths, i.e. according to curtate duration at death and age last birthday on 1st January falling in the policy year of death, instead of merely by curtate duration and age last birthday at death. No question of principle is involved, and it is sufficient for the student to understand the essential difference between the two methods. There is no published record of the modified policy-year method having been used.

Example 1.

In an investigation it is found that at any point of time the number of new entrants is the same at each point of age in the region of entry age x. The number of new entrants in 1930 at age x last birthday at time r after the beginning of the calendar year was proportional to $\frac{2}{3}+r^2$. In 1931 and during the first half of 1932, the rate of flow of new entrants was the same as at the corresponding date in 1930, but in the middle of 1932 the rate was suddenly reduced by one-half and continued on this basis during the remainder of 1932 and in 1933–35 inclusive.

Considering new entrants only and neglecting exits, find the error

introduced at duration 0 by using the census method separately for each of the years 1931–35.

If N is the total number of new entrants in 1930 at age x last birthday, then the number entering in the interval r to $r+\Delta r$, where Δr is small, will be $N(\frac{2}{3}+r^2)\Delta r$, since $\int_0^1 (\frac{2}{3}+r^2)\,dr=1$. This expression also gives the number of entrants at age $x-1$ last birthday in view of the statement at the beginning of the question, and the corresponding numbers in 1931 and the first half of 1932. In the middle of 1932 the rate of flow was suddenly reduced by one-half and, for the remainder of 1932 and the whole of 1933, 1934 and 1935, the number of new entrants between times r and $r+\Delta r$ measured from the beginning of the appropriate calendar year was therefore $\frac{1}{2}N\,(\frac{2}{3}+r^2)\,\Delta r$.

Let us examine the progression of ${}^sP_x^0$ in each of the years 1931–35, s being measured from the beginning of the appropriate year. ${}^sP_x^0$ on any date is the same as the number of new entrants at age x last birthday during the year ended on that date, for although some of these entrants will have reached age $x+1$, they will have been replaced by lives who entered at $x-1$ last birthday.

Hence if r is the time of entry measured from the beginning of the year of entry, ${}^sP_x^0$ in 1931 is made up of $\int_s^1 N(\frac{2}{3}+r^2)\,dr$ from entrants in 1930 and $\int_0^s N(\frac{2}{3}+r^2)\,dr$ from entrants in 1931. We have, therefore,

$$\begin{aligned} {}^sP_x^0 &= \int_s^1 N(\tfrac{2}{3}+r^2)\,dr + \int_0^s N(\tfrac{2}{3}+r^2)\,dr \\ &= N \end{aligned}$$

and

$$E_{[x]}^c = \int_0^1 N\,ds = N.$$

Since ${}^0P_x^0 = {}^1P_x^0 = N$, the census method gives the correct value of $E_{[x]}^c$, a result which could have been deduced from the fact that condition (*b*) (ii) of paragraph 6 is satisfied.

In 1932, ${}^sP_x^0 = N$ during the first six months, while, during the latter half of the year,

$${}^sP_x^0 = \int_s^1 N(\tfrac{2}{3}+r^2)\,dr + \int_0^{\frac{1}{2}} N(\tfrac{2}{3}+r^2)\,dr + \int_{\frac{1}{2}}^s \tfrac{1}{2}N(\tfrac{2}{3}+r^2)\,dr,$$

the first integral relating to entrants after time s in 1931, the second to entrants in the first half of 1932 and the third to entrants between times $\frac{1}{2}$ and s in 1932 after the fall in the flow of new entrants had occurred.

Therefore $\quad {}^sP_x^0 = N - \int_{\frac{1}{2}}^{s} \tfrac{1}{2}N(\tfrac{2}{3}+r^2)\,dr = N\left(\frac{19}{16}-\frac{s}{3}-\frac{s^3}{6}\right)$

and $\quad E_{[x]}^c = \int_0^{\frac{1}{2}} N\,ds + \int_{\frac{1}{2}}^{1} N\left(\frac{19}{16}-\frac{s}{3}-\frac{s^3}{6}\right)ds$

$$= \cdot 930N.$$

Now ${}^0P_x^0 = N$ and ${}^1P_x^0 = \frac{11}{16}N$, so that by the census method

$$E_{[x]}^c = \tfrac{27}{32}N = \cdot 844N,$$

involving an error of about 9 per cent of the true value.

A fairly substantial error was to be expected, for there is a considerable difference both in the number of lives under review at the beginning and end of the year and in the distribution of their entry dates.

The methods used to test the census method for the year 1932 can also be employed for 1933, and it can be shown that the correct value of $E_{[x]}^c$ is $\cdot 549N$ and the value by the census method is $\cdot 594N$. The percentage error is about 8.

In 1934 and 1935, however, the effect of the change in the flow of new entrants will have been stabilized. ${}^sP_x^0$ will be $\frac{1}{2}N$ throughout each year and this will also be the value of $E_{[x]}^c$. The census method gives the same value.

It is apparent that in the individual years 1932 and 1933 the census method is not suitable for duration 0 without adjustment.

Example 2.

Use the data of Example 1 to find the error introduced by the census method when the investigation covers the whole period 1931–35.

To test the suitability of the census method, we need only examine the lives under review at the beginning of 1931 and at the end of 1935. At curtate duration 0, we should find N lives aged x last birthday on the former date and $\frac{1}{2}N$ on the latter date, the distribution of policy anniversaries being the same in each case. Then

$${}^s\overline{P}_x^0 = \int_s^1 N(\tfrac{2}{3}+r^2)\,dr + \int_0^s \tfrac{1}{2}N(\tfrac{2}{3}+r^2)\,dr.$$

The first term relates to lives in 1931 who at time s had not reached their policy anniversaries in that year (i.e. entrants in 1930), and the second term relates to entrants in 1935. ${}^s\overline{P}_x^0$ reduces to $N\left(1-\frac{s}{3}-\frac{s^3}{6}\right)$ and $\int_0^1 {}^s\overline{P}_x^0\,ds$ to $\cdot 792N$.

The terms in formula (4) approximating to this part of the exposed to risk are $\frac{1}{2}({}^0P_x^0 + {}^5P_x^0) = \cdot 75N$. Hence the error by the census method is $\cdot 042N$.

The remaining terms in formula (4) consist of the census populations at age x last birthday, curtate duration 0, at the beginning of the years 1932–35 inclusive. We see from Example 1 that their sum is

$$N(1 + \tfrac{11}{16} + \tfrac{1}{2} + \tfrac{1}{2}) \simeq 2\cdot 687N.$$

Hence the value of $E_{[x]}^c$ by the census method is $2\cdot 687N + \cdot 75N = 3\cdot 437N$. As we have already seen, the error is $\cdot 042N$, so that the correct value of $E_{[x]}^{c1}$ is $3\cdot 479N$.

We can check the latter value and incidentally confirm to some extent the correctness of the argument in paragraph 8, by summing the values of $E_{[x]}^c$ for the individual calendar years which we calculated in Example 1:

$$E_{[x]}^c = N(1 + \cdot 930 + \cdot 549 + \cdot 5 + \cdot 5) = 3\cdot 479N.$$

In this example, the use of the census method without adjustment causes an error of a little over 1 per cent.

At duration 1, the effect of the fall in the number of new entrants would not be felt until a year later, i.e. the middle of 1933. Over the five years 1931–35 the error would be the same as at duration 0, provided that the number and distribution of new entrants in 1929 were the same as in 1930 and 1931. This is also true of duration 2, subject to a similar proviso regarding 1928 entrants.

$E_{[x]+3}^c$, however, is unaffected until 1935 and ${}^5P_{x+3}^3$ accordingly has a different distribution of policy anniversaries from ${}^0P_{x+3}^3$ instead of merely differing in size as did the corresponding functions at earlier durations. The error in ${}^s\overline{P}_{x+3}^3$ will be the same as for ${}^sP_x^0$ in 1932, i.e. $\cdot 086N$ (Example 1), since the two functions are identical for all values of s between 0 and 1. The correct value of $E_{[x]+3}^c$ for the five years is $4N + \cdot 930N = 4\cdot 930N$, and the percentage error when the census method is used is therefore about 2.

It should be noted that, as before, we have neglected all exits. This assumption does not vitiate the general argument.

The method adopted for testing the suitability of the data takes account of fluctuations in the flow of new entrants *before* the period of the investigation. If in our example the flow of new entrants had changed in 1929, this would affect the progression of ${}^sP_{x+1}^1$ in 1930, ${}^sP_{x+2}^2$ in 1931, etc., and would introduce errors in the values of $E_{[x]+1}^c$, $E_{[x]+2}^c$, etc., by the census method. The existence of these errors would, however, be revealed by comparing ${}^0P_{x+1}^1$ with ${}^5P_{x+1}^1$, ${}^0P_{x+2}^2$ with ${}^5P_{x+2}^2$ and so on.

When more than one variation in the flow of new entrants has occurred, our test takes account of the net effect only. If, for example, the rate of

flow of new entrants was doubled in the middle of 1929 and fell to its earlier level in the middle of 1932, ${}^{0}P^{2}_{x+2}$ and ${}^{5}P^{2}_{x+2}$, in the absence of any other disturbing factors, would be equal and would have the same distribution of policy anniversaries. We should, therefore, conclude that $E^{c}_{[x]+2}$ by the census method would be sufficiently accurate without adjustment.

The reader can check the correctness of this conclusion in the case of our example by working out the values of $E^{c}_{[x]+2}$ for individual calendar years. It will be found that, if $\frac{1}{2}N$ is the annual number of new entrants in 1928 and in 1932–33, the exposure for the five years is

$$N(\cdot 570+\cdot 951+1+\cdot 930+\cdot 549)=4N.$$

By formula (4), the value of $E^{c}_{[x]+2}$ is

$$N\left(\tfrac{1}{2}\times\tfrac{1}{2}+\tfrac{13}{16}+1+1+\tfrac{11}{16}+\tfrac{1}{2}\times\tfrac{1}{2}\right)=4N.$$

CHAPTER XIII

THE GENERAL THEORY OF SELECTION

1. Definition of selection.

In actuarial literature the word *selection* is generally used in a special sense. It may be defined as the operation of any variable factor, other than age, which tends to influence rates of mortality. Sex, occupation and nationality are examples of factors which exercise selective influences. The effect of such factors is to make the data heterogeneous and, as explained in Chapter I, we have to decide what subdivisions are necessary to secure a sufficiently high degree of homogeneity.

It is convenient to classify the various types of selection into four main groups:

(i) temporary initial selection,
(ii) class selection,
(iii) time (or generation) selection,
(iv) spurious selection.

Of these, group (i) has been fully discussed in Chapter XI and the methods of constructing the required select tables explained. In the light of the general definition of selection given above, however, one point requires clarification.

Reference was made in (XI, 2) to the fact that proposers accepted at ordinary rates by a life office are a selected class. In the actuarial sense, this use of the word 'selected' is strictly correct only if we are concerned with the proposers who were not accepted at ordinary rates as well as those who were, in which case the effect of the selective influence will be manifest in the lower mortality experienced by the latter group. When constructing a select table, however, we are dealing only with those proposers who were accepted at ordinary rates, and the particular type of selection which concerns us is that which makes the rates of mortality vary according to duration since entry. If the effect of initial selection were constant instead of variable, the rates would not depend on

duration and it would not be necessary to construct a select table —in other words, the class of proposers accepted at ordinary rates would not reveal any selection due to the production of evidence of health at entry, for all the lives at a given attained age would be equally affected and there would be no variable factor as required by the definition.

2. Class selection.

As the name implies, *class selection* is caused by the existence of certain attributes, each of which leads to the division of lives into a number of mutually exclusive classes according to the different variations of the attribute. If the mortality of the lives varies in the different classes, selection as defined at the beginning of the chapter clearly exists.

Some selective attributes, e.g. sex, are constant throughout life though their effect on rates of mortality may vary from age to age. Other attributes, e.g. occupation and place of residence, may vary throughout an individual lifetime and obviously lend themselves to a very wide variety of subdivisions and groupings. The problems of class selection are therefore of considerable complexity.

The difference between temporary initial selection and class selection is not easily expressed in words. Temporary initial selection is the result of an incident on a particular occasion, e.g. acceptance for assurance as a first-class life, whose influence on the mortality of the lives affected varies with the period elapsed since its occurrence. Class selection, on the other hand, is the result of continuing circumstances whereby lives are divided into classes each of which has its own characteristics influencing mortality rates. The effects of the two are not always readily distinguishable and in cases where class selection varies throughout life an element of temporary selection may be introduced. If, for example, we examined the mortality rates of a group of lives who had formerly lived in a tropical climate and had retired to Great Britain, and compared them with the mortality rates of a group of lives similarly constituted as regards sex, age and occupation who had always lived in Great Britain, we should probably find that the difference in mortality would diminish as the number of years since retirement increased

and might eventually disappear altogether. This phenomenon would be analogous to temporary initial selection, the incident being the classification of the lives on retirement according to whether or not they had been employed abroad.

It is important to realize that the methods of dealing with class selection and temporary initial selection are very similar. For the former, we subdivide the data into separate classes until a sufficient degree of homogeneity has been attained (i.e. until the effects of class selection have been for practical purposes eliminated) and construct separate tables for each class. For the latter, we construct a select table, which is of course merely a convenient way of setting out a number of different tables applicable to different entry ages which could be tabulated separately if we so wished. Since the effect of initial selection wears off as the duration increases, we can eventually combine the data to form an ultimate table and so obtain a single table in a convenient form. The same process could be followed for different classes of lives, if it were found that after a certain age the difference in the rates ceased to be of practical significance.

3. Class selection: the problem of heterogeneity.

When class selection exists, the data are heterogeneous. As explained in Chapter I, this reduces the value of the resulting rates and may for some purposes render them quite worthless. To overcome this defect, we must subdivide the data until each group is homogeneous, although in doing so there is a risk of so reducing the numbers in some or all of the groups that the results are unreliable. How far homogeneity should be sacrificed in order to avoid too drastic a reduction in the amount of data in each group is a problem which seldom admits of an obvious solution and which has given rise to much controversy in some of the best-known investigations.

One of the main considerations to bear in mind is the object of the investigation. Broadly speaking, mortality rates are required either for purposes of comparison or as a guide to the future. An example of the former is an investigation into the mortality of lives engaged in different occupations to test the effect of occupation on mortality. For this purpose the emphasis must be on homogeneity,

and we must examine the data to see whether differences in the mortality of two occupational groups might not be due to selective influences other than occupation. It would not, for instance, be satisfactory to include both male and female lives in the same investigation, since sex has an appreciable effect on mortality, and the proportion of males to females varies considerably between one occupation and another.

The use of past experience as a guide to the future is typified by the important problem of investigating the experience of life offices to obtain rates for calculating premiums and reserve values. Proposers for life assurance are drawn from different classes of society, different occupations and different parts of the country and so may form an exceedingly heterogeneous body of data even after medical evidence has excluded those with serious physical impairments. It would, however, be highly inconvenient for an office to publish a large number of different tables of premiums applicable to the various districts and occupations, even if the data were sufficient to permit of their being subdivided into so many different groups, and it is therefore fortunate that the nature of the business allows offices to proceed on broad general lines. Life assurance is based on the principle of averages, and so long as the total future premiums received are likely to be adequate to meet the total future claims the office will remain solvent. From this standpoint heterogeneity in the data is not important, provided that the lives assured are made up of a similar mixture to those who contributed to the experience on which the premium rates were based. In general, there is no reason to expect any important change in the mixture. It is therefore customary to make only a few subdivisions in the data and to exclude from the main body lives subject to considerable extra risks on account of occupation, physical impairment or place of residence.

In view of the above argument, it may be asked why any allowance at all need be made for class selection. There are two reasons. First, in the interests of equity it is undesirable that the same rate of premium should be charged for two classes which are likely to experience widely different rates of mortality. Secondly, there is the risk of options being exercised against the office by the insuring

public. The meaning of this can be explained by an example. Suppose that an office which had previously charged an extra premium for publicans decided to accept them in future on ordinary terms and increased its normal premium rates so that they were based on the mortality experienced in the past by its policyholders including publicans. The natural result of this change would be an increased flow of proposals from publicans and a reduction in the number of proposals from first-class lives who would be able to obtain better terms elsewhere. Other things being equal, the increased proportion of publicans in the lives assured would increase the rates of mortality experienced by the policyholders as a whole. To meet this, the office would have to increase its premium rates again, as a result of which first-class lives would be further discouraged from making proposals and the general standard would fall still lower. This argument would not apply to the same extent if all offices changed their practice and increased their rates simultaneously, and in that event the only argument would be on the grounds of equity.

4. Time selection.

It is well known that mortality rates in this country have decreased with the passage of time. This improvement, which is the result of advances in medical science, improved hygiene, general betterment of living conditions and other factors, has continued for many years and has often been sufficiently rapid to produce a marked effect within a comparatively short period. Time is therefore a selective influence within the meaning imposed by the definition. We should probably find, for example, that a group of lives aged x in 1935 would exhibit a lighter rate of mortality at that age than the rate experienced at the same age by a similar group who were aged x in 1920. An experience covering the period 1920–35 would therefore be heterogeneous as a result of *time selection* and would require to be subdivided into shorter periods covering a few calendar years. It was one of the major criticisms of the British Offices' Experience (which covered the period 1863–93) that, in order to make the data as extensive as possible, the period of the investigation was so prolonged that time selection had a serious effect. There

has accordingly been a tendency in more recent years to confine investigations to shorter periods.

The existence of time selection puts new difficulties in the way of using past experience as a guide to the future. We cannot, for example, assume that the rates of mortality experienced by a group of lives in the years 1924–29 will apply to another group of lives in the years 1944–49, even if the constitutions of the two groups are similar. To overcome this difficulty there has been devised the process known as projection, which we shall consider in a later chapter. In the meantime, we note the fact that only if the period of the investigation is kept short will the data be homogeneous so far as time selection is concerned.

The time factor may enter into our calculations in another way by producing changes in the incidence of other forms of selection. For example, the effect of temporary initial selection on the mortality of lives assured is by no means constant. The growth of medical knowledge and the introduction of new methods of investigation and diagnosis have increased the efficacy of medical selection and this in turn has increased the effect of temporary initial selection on the mortality of the selected lives. Class selection may be similarly influenced by time. For example, a change in general financial conditions may alter the type of lives who effect immediate annuities and such variations must be borne in mind in any investigation. These are particular cases of the general problem of spurious selection which we discuss below.

5. Spurious selection: definition and examples.

In the particular case where the duration of temporary initial selection is distorted by other selective influences, we have the fourth type of selection referred to in paragraph 1. *Spurious selection* may be defined as the apparent existence of temporary initial selection to a greater or less extent than is in accordance with the facts. This is not really selection at all, but merely a statistical phenomenon caused by heterogeneous data. There are two types of spurious selection: (i) positive, in which the degree of temporary initial selection is over-emphasized or in which an appearance of selection is created where none is really present, and

(ii) negative, in which the presence of temporary initial selection is wholly or partly concealed. The following examples may make these points clear.

Example 1.

The following table shows the experience of males and females at attained age $x+t$:

	$E_{[x]+t}$	$\theta_{[x]+t}$	$q_{[x]+t}$	$E_{[x-1]+t+1}$	$\theta_{[x-1]+t+1}$	$q_{[x-1]+t+1}$
Males	1000	50	·050	2000	100	·050
Females	2000	80	·040	1000	40	·040
	3000	130	·043	3000	140	·047

It will be seen that in the separate experiences temporary initial selection has worn off by duration t so that $q_{[x]+t}=q_{[x-1]+t+1}$, but in the combined experience the appearance of temporary initial selection has been caused by combining in varying proportions at durations t and $t+1$ two classes of lives experiencing different rates of mortality.

Example 2.

Calendar years of exposure	$E_{[x]+t}$	$\theta_{[x]+t}$	$q_{[x]+t}$	$E_{[x-1]+t+1}$	$\theta_{[x-1]+t+1}$	$q_{[x-1]+t+1}$
1920–25	1000	40	·040	2000	80	·040
1925–30	1200	43	·036	1000	36	·036
1930–35	1800	58	·032	1000	32	·032
1920–35	4000	141	·035	4000	148	·037

Heterogeneity due to time selection may have a similar effect, as the above example shows. The figures 1920–25 in the first column signify that the corresponding years of exposure are those falling between the policy anniversaries in 1920 and 1925.

Examples 1 and 2 both exhibit positive spurious selection.

Example 3.

	$E_{[x]+t}$	$\theta_{[x]+t}$	$q_{[x]+t}$	$E_{[x-1]+t+1}$	$\theta_{[x-1]+t+1}$	$q_{[x-1]+t+1}$
Males	2000	108	·054	1000	58	·058
Females	1000	42	·042	2000	92	·046
Total	3000	150	·050	3000	150	·050

In this example, temporary initial selection is present at duration t

in both classes; when they are combined, the difference in the proportion of male to female lives at durations t and $t+1$ hides the existence of this selection.

6. Indications of spurious selection.

Spurious selection, like initial temporary selection, is most commonly encountered in life office experiences. It would help us to recognize the existence of spurious selection if we knew how long initial temporary selection would normally last, for we should then suspect the presence of positive spurious selection if the select period seemed exceptionally long and of negative spurious selection if the reverse were the case. Unfortunately, no such guidance is available, for it has not been found possible in past experiences to make a clear distinction between the effects of temporary and spurious selection. In the $O^{[M]}$ table, based on the experience of certain assured lives in Great Britain during the years 1863–93, the select period was taken as ten years, but no account was taken of the possibility of spurious selection, which had not then received so much attention as in recent years; it is now considered likely that the effect of initial selection was overstated. In the A 1924–29 experience the possibility of spurious selection was fully appreciated, but the problem of measuring its effect was not solved and, in fixing the select period as three years, it was realized that the effect of initial temporary selection might not have been exhausted by the end of that period.

As regards life office annuitants, a five-year select period was adopted in the 1863–93 experience, while in the 1900–20 experience the chosen select period was as short as one year. Both in this experience and in the 1924–29 experience of assured lives, a short select period was adopted partly on the grounds of convenience.

Clearly, little help in detecting spurious selection can be obtained from this source. A more obvious guide to the presence of spurious selection is sometimes given by the failure of the select rates to tend towards the ultimate. For example, select rates of mortality might be obtained on the basis of a three-year select table, and it might be found that they did not run into the ultimate rates excluding the first three years. This would suggest that temporary initial selection

lasted for more than three years. If, however, it were found that with a five-year or seven-year select period we were not apparently much nearer to effecting a smooth junction between the select and the ultimate rates, there would be strong evidence of the existence of spurious selection. The following numerical example illustrates this.

Example 4.

x	$q_{[x-2]+2}$	$q_x^{(3)}$	$q_{[x-4]+4}$	$q_x^{(5)}$	$q_{[x-6]+6}$	$q_x^{(7)}$
30	·00250	·00290	·00260	·00300	·00270	·00310
35	·00300	·00345	·00310	·00355	·00320	·00365
40	·00400	·00450	·00415	·00465	·00425	·00475

The symbol $q_x^{(t)}$ represents the ultimate rate of mortality excluding the first t years' exposure after entry. It will be seen that $q_{[x-t]+t}$ increases with t and $q_x^{(t+1)}$ increases at the same rate, so that there is no apparent tendency for the select rates to run into the ultimate.

7. Causes of spurious selection: (i) variation of mortality with year of entry.

Results of the type illustrated in Example 4 above were obtained in the 1900–20 experience of life office annuitants, and the problem of spurious selection therefore received particular attention. Several possible explanations of its presence were put forward, in particular the effect of variations in the rate of mortality according to calendar year of entry. It was pointed out that the shorter the duration in force the later was the average calendar year of entry and that, if there were factors causing an improvement in mortality according to year of entry, this would result in spurious selection. A possible factor answering this description would be a progressive increase in the proportion of new entrants drawn from classes experiencing light mortality and a corresponding decrease in the proportion drawn from classes showing heavy mortality. Example 1 (p. 137) illustrates this kind of spurious selection. The data for entry age $x-1$ duration $t+1$ would on the average arise from lives who entered about a year sooner than the lives contributing to $E_{[x]+t}$. Within that year there must therefore have been a marked increase in the

proportion of female entrants and a corresponding reduction in the proportion of males. The example is, of course, an extreme one, and it is not likely that in practice a switch-over would take place as quickly.

Whatever may be the primary cause, there is no doubt that spurious selection is bound to occur if the net effect of the factors (other than temporary initial selection) influencing mortality is to make the rates depend on the calendar year of entry. The following example shows this effect clearly.

Example 5.

Let us suppose that we are investigating the mortality experience of a group of lives from their policy anniversaries in 1936 to those in 1939. The data relate to durations 5 and 6, so that the calendar years of entry with which we are concerned are 1930–33.

Year of exposure	Calendar year of entry	$E_{[x]+5}$	$\theta_{[x]+5}$	$q_{[x]+5}$
1936–37	1931	1000	36	·0036
1937–38	1932	1000	32	·0032
1938–39	1933	1000	29	·0029
1936–39		3000	97	·0032
Year of exposure	**Calendar year of entry**	$E_{[x-1]+6}$	$\theta_{[x-1]+6}$	$q_{[x-1]+6}$
1936–37	1930	1000	40	·0040
1937–38	1931	1000	36	·0036
1938–39	1932	1000	32	·0032
1936–39		3000	108	·0036

The figures 1936–37 in the first column signify that the exposure extended from the policy anniversaries in 1936 to those in 1937. The years 1931–33 are the only years of entry giving rise to exposure at duration 5 between the policy anniversaries in 1936 and 1939 and similarly for the years of entry 1930–32 and duration 6.

The rates of mortality for individual years of entry do not show any evidence of temporary initial selection, being the same at durations 5 and 6, but over the experience as a whole the appearance of initial temporary selection is created.

8. In Example 2, spurious selection would not have arisen but for the fact that the proportions of the exposed to risk at durations t and $t+1$ respectively were not the same in each of the three groups of calendar years. This phenomenon would almost certainly be caused by fluctuations in the number of new entrants at ages $x-1$ and x. In this example these fluctuations must have been very marked to have produced the values of $E_{[x]+t}$ and $E_{[x-1]+t+1}$.

In Example 5, however, the proportions at durations 5 and 6 respectively are the same in each year of exposure, but spurious selection is nevertheless apparent. Clearly, therefore, the dependence of mortality rates on calendar year of entry has greater potentialities as a cause of spurious selection than the existence of time selection as described in paragraph 4. The former phenomenon is sometimes regarded as a form of time selection, but it is better to limit this designation to the influence whereby q_x depends on the calendar year in which age x is attained—or, what is in effect the same thing, the calendar year of birth.

It is important to realize that, if rates of mortality are such that q_x depends solely on calendar year of entry, a rather paradoxical situation arises. For example, two groups of lives entering in 1900 at ages 20 and 40 respectively would experience the same rate of mortality at age 70, although the first group would not attain this age until 1950 whereas the second reached age 70 in 1930. Again, two groups attaining age 70 in 1940 would experience different rates of mortality at that age if they had entered at different times, even although temporary initial selection had passed off.

We shall consider this problem again when we are investigating the subject of projection.

9. Causes of spurious selection: (ii) use of group rates of mortality.

There is one other cause of spurious selection to which reference must be made. This cause is liable to operate when comparisons are made of group rates of mortality, i.e. rates based on the data for a group of ages instead of a single age. The object of such grouping is to save time in carrying out preliminary investigations into the effect of selection, etc., and to obtain a larger amount of data for each

unit of age so as to reduce random errors. If the proportions of the exposed to risk at the various ages in a group vary appreciably between one class and another, a comparison of the rates for the two classes may be misleading. In particular, if we are comparing groups of lives based on select and ultimate data, such a variation will result in spurious selection. This state of affairs is very liable to occur in practice, for the building up of the ultimate data from the select automatically leads to different weighting of the select and ultimate rates. Over the lower range of ages covered by the experience the number of new entrants will usually change fairly slowly as age advances. The same is therefore true of the select exposed to risk. The ultimate data, on the other hand, are being continually increased by the transfer of lives from the select data and, as the number of exits at the younger ages is comparatively small, the net effect at these ages is usually a steady increase in the value of E_x as x increases. This can be seen by examining the exposed to risk in the A 1924–29 experience or in the 1900–20 annuitants' experience. A simple illustration may, however, be helpful.

Suppose that there were no new entrants under age x and that there were N entrants per annum at each age from x to $x+9$ inclusive. The exposed to risk in an experience extending from the policy anniversaries in one year to those in the following year would be as follows, if we were to neglect exits and assume a 3 years' select period.

Age at entry y	$E_{[y]}$	$E_{[y-1]+1}$	$E_{[y-2]+2}$	$E_y^{(3)}$
x	N			
$x+1$	N	N		
$x+2$	N	N	N	
$x+3$	N	N	N	N
$x+4$	N	N	N	$2N$
...	...	...	...	...
...	...	...	...	...
...	...	...	...	...
$x+9$	N	N	N	$7N$

It is clear that, for any group of ages, the average age is older for the ultimate than for the select data, and this will, of course,

create an appearance of temporary initial selection. There are other factors at the older ages influencing the relative trends of the select and ultimate data, and it is more difficult to say whether spurious selection is likely to be positive or negative.

When group rates of mortality are to be calculated, the data should always be examined for differences in the average ages of the select and ultimate exposed to risk. A rough allowance can then be made for the effect of such differences so as to avoid drawing erroneous conclusions as to the extent to which temporary initial selection is operating.

10. Multiple selection.

The following example illustrates the method of approach when several selective influences are operating on the data simultaneously.

Example 6.

Suppose that we are investigating the mortality of the staff of a group of factories situated in different parts of the country, employing clerical and manual workers of both sexes. Our knowledge of past experiences might lead to the hypothesis that the most powerful selective influences would be sex and nature of work, i.e. whether clerical or manual, and that a secondary influence would be geographical situation.

On this hypothesis, we would segregate the data for the sexes and subdivide according to nature of work, thus obtaining four groups for each of which we would set out the rates of mortality for quinquennial or decennial age groups (i.e. the data for ages 20–24, 25–29, etc., or for 20–29, etc., would be aggregated and a rate calculated for each age group instead of obtaining a separate rate for each individual age: if the latter course were followed, the resulting numbers would probably be too small to permit of any conclusions being drawn).

Table 1

	E	θ	q
Male clerks	25,000	400	·016
Female clerks	5,000	70	·014
Male operatives	10,000	200	·020
Female operatives	100,000	1800	·018

Table 1 is an example of the results which might be obtained for such an age group. From these figures, we would reach the tentative

conclusion that sex and nature of work had some influence on the mortality of the age group in question, since the rates for female clerks and operatives are less than those for male clerks and operatives respectively, while the rates for male and female clerks are less than those for male and female operatives respectively. The differences might, of course, be due to random errors, but for present purposes we shall neglect this possibility.

Let us now see what the result would have been if we had compared the mortality of males and females without subdividing according to occupation and vice versa.

Table 2

	E	θ	q
Males	35,000	600	·017
Females	105,000	1870	·018
Clerks	30,000	470	·016
Operatives	110,000	2000	·018

Table 2 suggests that the rate of mortality for males is less than that for females, but the more complete analysis we have already made shows that such a conclusion would be incorrect. The distortion is caused by the fact that there is a higher proportion of operatives among the females than among the males. The comparison of the mortality of clerks and operatives is also distorted by combining males and females, but the distortion is less pronounced.

This example illustrates the danger of basing comparisons on heterogeneous data.

Having arrived at certain tentative conclusions about the influence of sex and occupation, we would next consider the effect of geographical situation. Our previous knowledge would indicate what subdivisions should be suitable, and we shall suppose that the country has been divided into three areas for the purpose. This would involve twelve separate groups of data, i.e. area 1, male clerks; area 2, male clerks, etc. In order to test the effect of geographical situation, we should have to compare the rates for male clerks in each of the three areas, then the rates for male operatives and so on. If the data were sufficiently numerous, the comparison should give some indication of the effect of geographical situation on the mortality of the employees.

Should the influence of the last factor appear to be pronounced, we should have to reconsider the first two factors, i.e. sex and occupation, in order to see whether our conclusions had been falsified by operating with data heterogeneous as to geographical situation. This would involve a separate comparison of the rates for males and females respectively in

each of six groups, i.e. area 1, operatives; area 1, clerks, etc., and a similar comparison of the rates for clerks and operatives in each of six groups, e.g. area 1, males; area 1, females, etc. If the geographical distribution was roughly the same within each of the four groups, male clerks, female clerks, male operatives and female operatives, the original comparison neglecting geographical situation would not be falsified by the heterogeneity due to this factor.

We have assumed that the data are adequate to allow subdivision into twelve groups. It is possible, however, that in at least some of the groups the numbers would be so small as to be unreliable. In order to test the effect of geographical distribution, we might then have to combine the data into one group for each area irrespective of sex and occupation. We should then be faced with the problem of deciding how far the differences revealed by a comparison of the three groups were due to geographical influences and how far they were caused by the heterogeneity of the data. Alternatively, we might be able to operate with the twelve subdivisions if we increased the size of the age groups, e.g. from five to ten or even fifteen individual ages. This alternative would commend itself if the age distributions within the enlarged age groups were roughly the same in each of the twelve subdivisions.

Clearly it is not possible to lay down any rules for solving such problems. They demand from the actuary an intimate knowledge of the pitfalls which may be encountered and, in addition, considerable ingenuity. Much depends on the object of the investigation.

11. Methods of investigating selection.

By this stage the reader may be inclined to think that there are so many causes of selection as to make the construction of a table of mortality rates based on homogeneous data an almost impossible feat. It is certainly true that the task is one requiring the exercise of careful discrimination and balanced judgment on the part of the actuary. One of the chief difficulties which faces the student is the impossibility of laying down any general method of dealing with the problem, as every investigation has its own peculiar features. Careful study of some of the major investigations of the past will make the subject seem more real, although the accounts of these investigations are sometimes reticent about the groupings and regroupings of the data which were tried and discarded and deal only with those finally adopted.

As a further aid, the following summary is given of the steps which might be taken in the preliminary investigations before

calculating the final tables of rates. The assumption has been made that the rates are subject to temporary initial selection. It will be appreciated that this summary indicates only one order of approach and that different operators would no doubt proceed along different lines.

(*a*) In the light of previous knowledge of similar data, estimate roughly the length of the select period and aggregate the data for longer durations—referred to hereafter as the ultimate data—which should then be virtually undisturbed by the effect of temporary initial selection when other kinds of selection are investigated.

(*b*) Group the ultimate data in quinquennial age groups. A smaller range of ages might be included in each group if the data are very plentiful, while decennial age groups might have to be used in a small experience.

(*c*) Consider what types of class selection are likely to have operated—again in the light of past experience. Subdivide the data according to the variations of each factor in turn and examine the results, bearing in mind the danger that these may be misleading (see Example 6).

(*d*) Keeping in view the uses to which the finished tables will be put and the extent of the class influences revealed, decide what separate classifications are desirable.

(*e*) If the experience covers a period of more than, say, 10 years, test for time selection by dividing the ultimate data in each class into groups covering a few calendar years each and compare the rates of mortality. Should the differences be considerable, separate tables for the different periods may be desirable.

(*f*) For each class obtained by carrying out the subdivisions decided upon in (*d*) and (*e*) above, test the effect of initial temporary selection by comparing the group mortality rates at different durations. The factor mentioned in paragraph 9 should be borne in mind. Unless the data are very extensive, some or all of the classes may have to be aggregated for this test.

(*g*) If the select rates do not tend to run into the ultimate, there may have been influences at work as a result of which mortality depends on calendar year of entry. This may be due to a progressive change in the proportion of new entrants in different classes if

a number of classes have been aggregated. The spurious selection should then disappear when the rates for these classes are calculated separately.

(*h*) If spurious selection does not disappear, the data for a few consecutive calendar years of entry should be investigated, provided that the period covered by the experience is long enough to permit of this. In the 1900–20 experience of life office annuitants, the data from lives contributing to the select experience in the years 1900–7 were traced throughout the whole 20 years. If spurious selection should no longer be apparent in the experience confined to a few calendar years of entry, the actuary must decide, having regard to the use for which the tables are designed, whether separate tables based on these data are desirable or whether the data for all years of entry should be used and some artificial means devised of making the select rates run into the ultimate at the duration when temporary initial selection is estimated to have ceased. This aspect of the problem will be encountered again in later chapters.

A full investigation on the lines set out in (*a*) to (*h*) above would not be possible unless the data were very extensive.

CHAPTER XIV

THE CONSTRUCTION OF MORTALITY TABLES FROM LIFE OFFICE RECORDS

For the majority of actuaries the most important investigations from a practical point of view are those which seek to determine the rates of mortality that should be used in calculating life office premiums and reserve values. The principles and methods by which ordinary and industrial life assurance are transacted differ materially and it is convenient to consider the two separately. In this chapter we shall confine ourselves to ordinary business. The mortality of annuitant lives is also a distinct subject and is discussed separately in Chapters XVI and XVII.

1. The selection of lives assured.

Mortality rates obtained from general population statistics (e.g. the experience of the whole population published by the Registrar-General) are not suitable for ordinary life offices, since assured lives are, for several reasons, a selected class.

(i) Lives proposing for assurance are subjected by the office to a process of scrutiny, which usually includes either a medical examination or a fairly exhaustive enquiry into the proposer's medical history and state of health. Those who do not conform to the office's requirements are not accepted for life assurance. As we have already seen, this results in temporary initial selection, but it also gives rise to a more lasting effect due chiefly to the exclusion of lives who are subject to permanent extra risks, e.g. workers in hazardous occupations who would throughout the remainder of their lives exhibit heavier mortality than a similarly constituted group of lives in less dangerous occupations. Where the extra risks to which the lives are subject are not sufficiently serious to justify their exclusion from life assurance benefits, they may be accepted by the office on special terms. Such lives would probably be treated as a special class and would be excluded from any general investigation into mortality.

(ii) Assured lives are subject to a form of class selection. The majority are drawn from the upper and middle classes of society where higher standards of nutrition, housing, etc. result in lighter mortality than that experienced by the population as a whole.

(iii) Life assurance has a special appeal to the more provident members of the population, who are likely to live in a careful and well-regulated way and therefore to experience lighter mortality than the average.

(iv) A certain selection may be exercised against the office, since there may be a tendency for those who believe themselves to be under-average lives to realize more fully the need to make provision for their dependants. If this tendency is present, it will of course operate in the reverse direction to (i), (ii) and (iii), but the selection exercised by the office should be sufficiently strong to limit its effect severely.

In view of these influences it is clearly desirable that premium rates and reserve values should be based on the actual mortality experienced by lives assured: this has in fact been the practice, except in the early years of life assurance when no life office data were available.

It should be noted that in this paragraph we have been dealing with selective factors which make life office mortality differ from that of the general population. On the other hand, in Chapters XI and XIII, we were concerned with selective influences causing variations in the rates of mortality experienced by different groups of assured lives.

2. Requirements of a mortality table for office use.

Before considering various specific problems which arise in constructing mortality tables from the records of assured lives, it may be desirable to mention the principal requirements of a mortality table intended for office use.

For the calculation of non-profit premiums, the chief essential is that future rates of mortality should not be underestimated. Before adopting for its without profit premiums a standard table based on the experience of a number of offices, each office must be satisfied that its own experience has not been and is not likely to

be worse than the standard. For many years there has been a steady and general improvement in mortality, so that the practice of using a mortality table based on past experience as a measure of future mortality has produced an error on the side of safety. There has in consequence been a tendency to discount future improvements, a practice which involves serious risks.

The case of with profit premiums is different. The bonus loading will usually ensure that the rates of premium are at least adequate to provide the guaranteed benefits. The problem is, therefore, to fix premiums which will support the rate of bonus it is hoped to maintain and at the same time to achieve equity between the different entry ages and the different classes of assurance. With profit premium rates are altered less frequently than without profit rates, since any necessary adjustment can, if desired, be made by altering the rate of bonus for some or all classes of policy. In general, an office cannot hope to achieve equity except on broad lines and it will usually be content to use a standard table for calculating with profit premiums unless its own mortality shows major variations which may be expected to continue in future.

In calculating reserve values, safety is again the primary consideration. When choosing a table for this purpose, attention should be paid to the trend of the office's mortality rates age by age compared with those of the standard table, in view of the fact that the steepness of the mortality curve has an important influence on reserve values. Where there is any evidence that the mortality curve of the office's own experience is steeper than that of the standard table, an adequate reserve must be held in addition to the reserves brought out by the standard table.

3. Accuracy of the data.

Much of the value of a mortality investigation depends on the accuracy of the underlying information. In this respect, investigations into the mortality of assured lives are happily placed, for nearly all the material facts are carefully recorded and verified. Exact dates of birth and death are available and in most cases evidence of both will have been forthcoming. Dates of entry and

exit will usually be known. The labour of obtaining or employing exact information for every case may be prohibitive, but at least this information is available if we wish to use it. It is important to remember that, in an investigation of life office data, any approximate assumptions are made from choice and not of necessity and that, by sampling or otherwise, we can check their accuracy. There will, of course, always be cases for which exact information is not available or for which wrong information has been supplied (e.g. cases where age has not been admitted), but the errors from this source are unlikely to be numerous or to have any significant effect on the investigation.

One possible source of error is the failure of the parties interested to intimate to the office the death of the life assured. This is a rare occurrence, for there is a strong inducement to intimate deaths without delay. Even if the existence of a policy is overlooked by the executors at the time of death, it will usually come to light when the next premium renewal notice is received. The office may, however, lose contact with the life assured altogether when a policy has been converted into a paid-up assurance for a small sum, owing perhaps to non-payment of premiums; the failure of the executors to intimate death would not then be revealed until the life assured had apparently reached such an advanced age that the office was prompted to make enquiries. Such omissions cause the rates of mortality to be understated, for not only are the deaths understated at the actual age of death, but the exposed to risk is overstated at all subsequent ages for which the life is incorrectly included.

The latter factor is important, for its effect is cumulative and, at the advanced ages, the error in the exposed to risk may be important enough to affect materially the rates of mortality. There is, therefore, something to be said for excluding from the investigation all non-profit paid-up policies for small sums assured under which the life assured is aged 90 or over. The risk is less serious in the case of with profit policies, as the issue of the bonus notices will usually bring to light the previous death of the life assured. At ages below 90 the error is unlikely to be of any practical importance.

4. Unit of investigation.

Four different units have been suggested for investigations of life office mortality:

(i) *Lives.* Each life assured is included in the investigation once only, i.e. one unit is included in the exposed to risk for each full year during which the life is under observation and one unit in θ_x in the year of death. Clearly this method gives a true rate of mortality according to our definition.

(ii) *Policies.* Each policy contributes one unit to the exposed to risk for each full year that it is in force and one unit to θ_x in the year of death. A life who is assured under more than one policy is therefore included more than once in the exposed to risk at the appropriate ages and similarly at the age of death. By this method we obtain a rate of discontinuance of policies by death instead of a rate of mortality.

(iii) *Sums Assured.* Each policy contributes to the exposed to risk for each full year that it is in force and to θ_x at the age of death a number of units equal to the sum assured. This gives us a rate of claim or rate of payment of unit sum assured by death.

(iv) *Sums at risk.* Each policy contributes to the exposed to risk and to the deaths a number of units equal to the sum at risk, i.e. the difference between the sum assured and the reserve value, sometimes called the *death strain.* This gives what we may call a *rate of strain.*

Units (ii), (iii) and (iv) are the results of assigning different weights to the lives assured. We shall now consider what is the effect of these weighting processes on the results of the investigation.

5. Investigations based on policies.

The treatment of lives who are assured under more than one policy has been widely discussed. Whether the rate of discontinuance of policies by death is the same as the rate of mortality which would be obtained if all duplicates (i.e. all policies on lives who already have a policy in force) were eliminated depends on whether the mortality of lives having two or more policies is or is

not the same as that of lives having one policy each. There are several cases to be considered of which the following are the most important.

(*a*) *Two or more policies effected simultaneously.*

There is no reason to suppose that lives who elect to take several policies at the same time instead of a single policy are subject to any selective influence on that account. The distortion caused by taking policies instead of lives as the unit will therefore be confined to random errors. If, for example, there were one hundred lives, ninety-nine of whom had effected a single policy each at age 20 and the remaining one had effected twenty-one policies, each death would increase $q_{[20]}$ by ·01 if a unit were included for each life, but, if a unit were included for each policy, a single death might increase $q_{[20]}$ by either ·0083 or ·175. On the average, the effect would be the same in each case, since $\frac{99\times\cdot0083+1\times\cdot175}{100}=\cdot01$, but there is more scope for random errors in the latter case. If the number of policies on each life were the same, the two rates would, of course, be identical.

Clearly, there is no theoretical advantage in using policies instead of lives as the unit in case (*a*). It is quite a simple matter to eliminate simultaneous duplicates effected with the same office and this is usually done. It is more troublesome to eliminate the duplicates if two or more policies are effected simultaneously with different offices. The work involved in discovering them is not usually worth while.

(*b*) *Two or more policies effected at different times included in* (*i*) *select data,* (*ii*) *ultimate data,* (*iii*) *aggregate data.*

In case (i), the different policies are included in different groups of data and cannot correctly be regarded as duplicates. To eliminate any of the policies would merely cause a loss of valuable data and it is therefore usual to include a unit at each separate entry age.

When we come to (ii), the question is more open to argument. In the first place, it is possible that lives who effect a succession of policies are as a class of a different type from those who effect one policy only—they may, for example, be of a more provident type.

Secondly, a more important factor operates if temporary initial selection lasts longer than the period assumed in the select tables, for in that case the lives having more than one policy will probably be subject to lighter mortality through their having been selected more recently on the average. If, for example, 1000 lives who had entered at age 30 survived to age 35 and at that age 500 of them made new proposals and 450 were accepted at ordinary rates, then, assuming that temporary initial selection lasts for ten years and that a select period of only five years is assumed in the tables, the mortality rates based on policies will almost certainly be less between ages 40 and 45 than those based on lives. In the former case more weight is given to the survivors of the 450 lives who were reclassified as first class at age 35 than to the survivors of the remaining 550, some at least of whom were no longer first class lives at age 35.

A similar state of affairs would exist if the standard of selection exercised by the offices had been improving or if any other influence had been at work to produce improvement in mortality rates according to calendar year of entry.

The generally accepted view is that, as the ultimate table is a combination of a series of select tables, the data should be a combination of the data which would be used if the select tables were extended indefinitely, i.e. a unit should be included for each policy, except where two or more were effected at the same time. This course was followed for the ultimate sections of both the $O^{[M]}$ and the A 1924–29 tables. In the case of the latter, however, offices were asked to eliminate duplicates above age 80, if this could conveniently be done, in order to reduce the size of random errors, which are particularly troublesome when the data are few.

Turning now to case (iii), we can see from the numerical example under case (ii) that the presence of temporary initial selection will almost certainly lead to a difference in the rates based on lives and policies. Even if selection ceased after five years, the mortality rates of the 450 lives reselected at age 35 would be lower between that age and age 40 than the rates experienced by the remaining 550 lives. The problem as to which unit should be adopted is therefore more important for an aggregate than for an ultimate

table. On the one hand, a rate based on lives is a more stable quantity, since it does not depend on the number of duplicates, which is itself a variable quantity. On the other hand, it may be argued that a rate based on policies is more suitable for use by an office to calculate reserve values, since the unit for valuation purposes is the policy and not the life. In this connection it should be noted that aggregate tables from life office data are seldom used except for calculating reserve values.

In constructing the O^M table, extreme care was expended in eliminating duplicates. It is very doubtful, however, if the work involved was justified, even if it could be established that it is more correct to use lives than policies as the unit, for at best an aggregate table is not a particularly delicate instrument and the difference in the rates for lives and policies is unlikely to be of much practical significance. An aggregate table has not been constructed from the data of the continuous investigation which has been going on since 1924, but if it should ever be decided to construct another such table based on a large quantity of data, it is improbable that it would be considered worth while to eliminate duplicates, except possibly at ages where the data were few.

6. Investigations based on sums assured.

The argument in favour of using the sum assured as the unit is that mortality rates depend on the size of the policy and that, in particular, lives holding very large policies experience on the average heavier mortality than assured lives as a whole. Evidence in favour of the last statement has been adduced in America, but there is no reliable evidence applicable to Great Britain.

An investigation based on sums assured takes into account the financial effect of death and life offices are of course primarily concerned with this aspect of the matter. On the other hand, (*a*) if mortality does not depend on sum assured, there is no point in taking sums assured as the unit, as this would merely increase the liability to random errors, and (*b*) if mortality depends on sum assured to an appreciable extent, separate tables should be constructed for different ranges of sums assured, e.g. £500–£999, £1000–£5000 and over £5000—otherwise the rates of mortality

would only be applicable to an office having a similar distribution of sums assured to that of the whole experience, whereas in actual fact the distribution differs very considerably between one office and another.

The calculation of rates based on sums assured is not a method which has been used in any major investigation in this country, nor has it been thought necessary to subdivide the data according to size of policy. Certain American and Canadian tables have, however, been constructed on this basis, e.g. the American-Canadian Investigation 1915.

7. Investigations based on sums at risk.

This method takes into account the financial effects of deaths. In the case of an endowment assurance, the sum at risk decreases more rapidly than for a whole life assurance and the difference is most marked round about ages 55–65 when most of the endowment business is nearing maturity. If, therefore, there is an appreciable difference in the mortality of the two classes at these ages, rates for the combined classes will vary according to whether they were based on lives or sums at risk. It has been suggested that, as whole life mortality according to the latest experience of offices in this country is apparently heavier than endowment assurance mortality, the use of a combined table based on policies is dangerous because too little weight is given to the whole life data from the standpoint of financial effect.

The same objections apply to an investigation based on sums at risk as those advanced against the use of sums assured. Moreover, such an investigation would involve a prohibitive amount of work, as the weight applicable to each policy would vary according to the number of years in force as well as the class of policy.

An office which considers that its whole life mortality is heavier than that shown by the latest standard table can adjust for this by making a suitable addition to the premiums. This solution of the problem, though admittedly a rough and ready one, is far more satisfactory than the construction of a mortality table based on sums at risk.

8. Effect of withdrawals.

Rates of mortality may be influenced by the proportion of lives who withdraw at the age under review and at the ages immediately preceding it. This will be the case if the lives withdrawing experience different rates of mortality from those who remain. It is generally believed that policyholders tend to exercise an option against the office when they withdraw. It is clear that this will sometimes be the case, for a man on his death-bed would not be so foolish as to surrender his policy if he could possibly avoid it even although he were badly in need of the surrender value. Conversely, improvident people are much more inclined to surrender than their more thrifty neighbours and it is possible that the former class experiences heavier mortality. These two influences operate in different directions and it is difficult to say which is the more powerful. Statistical evidence of the effect of withdrawal is not easily obtained, for lives withdrawing cease to be under observation and their mortality experience is not available.

In general, the effect of withdrawals on the rates of mortality experienced by assured lives is unlikely to be of any practical significance, but exceptions may occur when some special factor is at work to make the number of withdrawals unusually large. For example, a considerable proportion of the lives assured under children's deferred assurances elect to take the cash option on attaining age 21 or age 25 as the case may be. It is therefore advisable, when investigating the mortality of such lives, to deal separately with commencing ages 21 and 25 until attained age 30 or thereabouts, for the mortality rates for a few years after the commencing date may exhibit reversed temporary selection, i.e. they may increase less rapidly than would normally be the case or may even decrease over a range of a few ages.

CHAPTER XV

HISTORY OF MORTALITY TABLES OF ASSURED LIVES

1. In the last chapter we discussed some of the points to be considered in investigating the mortality of lives assured. To illustrate how these problems have been dealt with in practice and to give some account of the historical development of the subject, we shall now describe briefly the principal mortality tables based on life office data in Great Britain. For the sake of completeness we include two tables prepared in the early days of life assurance from general population statistics, which were extensively used by the offices for lack of more suitable tables based on their own records.

It must be emphasized that the descriptions which follow are not intended to be complete. For full information reference should be made to the official reports on the various tables issued when they were published.

2. Northampton Table.

This table, which was published in 1783, was the first to be used extensively for life assurance purposes. It was based on the deaths in the years 1735–80 in a parish comprising the greater part of the town of Northampton, a record having been kept of the ages at death. In the absence of any corresponding population statistics, the exposed to risk at age x was obtained by summing the deaths at and above this age. In the life table, $l_x = d_x + d_{x+1} + \dots$, but the corresponding relation is not correct for recorded deaths unless the population at each age has remained stationary for many years past. This was not the case in the period covered by the Northampton data, as the birth-rate was increasing and in consequence the rates of mortality were overstated, especially at the younger ages.

The author of the table, Dr Price, was one of the pioneers in the construction of life tables and a strong advocate of a system of registration which would produce adequate data. The Northampton table was a small part of his work and it is unfortunate that, as a

result of the adoption of this table by a number of life offices, he is best known as the author of a table which is faulty in construction. The use of the Northampton table resulted in the offices making large profits on their life assurance business; severe losses were however sustained on annuity contracts.

3. Carlisle Table.

General population statistics were again used for the Carlisle table. This table was published in 1815, the data being obtained from the records of two parishes in Carlisle. Censuses were taken in the years 1780 and 1787, lives being classified according to sex and, in the first census, according to age also. The deaths were those for the years 1779–87. The data from the second census were divided into age groups by reference to the age distribution according to the first census and the mean populations and deaths were then used to calculate rates of mortality by the census method.

The data for male and female lives were combined in calculating the rates of mortality. From the standpoint of the life offices this was a disadvantage, as female lives constituted more than half the total and the percentage of these lives among proposers for assurance was comparatively small.

The Carlisle table has been extensively used and, owing to the large variety of monetary functions calculated therefrom, it is still occasionally employed for valuing reversionary interests when contingent benefits are involved.

4. Morgan's Equitable Table.

This table, published in 1834 and based on the experience of the Equitable Life Assurance Society, was the first to be constructed entirely from life assurance records. It is also the first recorded instance of tracing the exposed to risk through calendar years. The assumption was made that the age next birthday at entry was attained on the following 1st January. The tabulation of the data was on different lines from those which would be adopted now, but the method of obtaining the exposed to risk was sound.

It is also noteworthy that the author realized the effect of

duration on rates of mortality, although no select tables were published—the data could hardly have been adequate for this purpose.

5. Seventeen Offices' Table.

The experience of a group of British offices for the period from their foundation up to 31st December 1837 was used for this table which was published in 1843. Returns were made by the contributing offices in schedule form showing, *inter alia*, the sex of each life and whether resident in town or country. The mortality of female lives was found to be heavier than that of male lives at ages up to 50 but lighter from ages 50–70. The calendar year method was again used.

It was pointed out by the investigation committee that the average duration of the policies included in the experience was only about eight years, owing to the fact that most of the contributing offices had been in existence for a comparatively short time and that the high proportion of select lives could not be expected to continue in future.

6. Institute of Actuaries Tables (H^M and H^F).

Twenty English and Scottish offices contributed data to this experience, the results of which were published in 1869. All the available data up to 31st December 1862 were included, so that the experience covered a wide range of years and was certainly not homogeneous as regards time. Cards instead of schedules were returned by the offices.

Once again the calendar year method was used, the new entrants at age x next birthday being assumed to enter in the middle of the calendar year and to attain age x on the following 1st January. The age at the end of the calendar year of exit was taken as the age next birthday at entry, plus the calendar year of exit minus the calendar year of entry, and withdrawals were assumed to pass out of observation in the middle of the calendar year. The formulae used were therefore

$$E_x = E_{x-1} + \tfrac{1}{2}(n_x + n_{x+1}) - \tfrac{1}{2}(w_x + w_{x+1}) - e_x - \theta_x \quad \text{and} \quad q_x = \frac{\theta_{x+1}}{E_x},$$

where the symbols have the same significance as in previous

chapters and x is in each case the assumed age at the end of the calendar year in which the event occurred. There is no term in b_x since all policies were included from their inception.

q_x was taken to be the rate of mortality at exact age x, but apparently no attempt was made at the time to investigate the correctness of this assumption, nor was the possibility taken into account that the average date of entry might not have been the middle of the calendar year. Subsequently it was discovered that the average length of exposure in the calendar year of entry was less than half a year.

In tabulating the data, the figures for each age at entry were recorded separately so that select tables might be prepared, but no such tables were in fact constructed until later. Separate tables were however constructed for males and females, excluding under-average lives in each case; these were designated the H^M (healthy males) and H^F (healthy females) tables. The experience is now usually known as the H^M experience. The $H^{M(5)}$ table, which excluded the calendar year of entry and the four succeeding calendar years (about $4\frac{1}{2}$ years of exposure), was also constructed. The H^M and $H^{M(5)}$ tables have been widely used for the calculation of premiums and reserve values and are still employed for the latter purpose by some offices.

7. H^M Experience: select tables.

Prior to the publication of the H^M table, several papers had already been written on the subject of selection. In 1876, George King submitted a paper to the Institute of Actuaries in which he emphasized the importance of this factor and gave the results of an investigation into the select rates of mortality according to the H^M data. These data were not in the most suitable form for the purpose, as the movements were not recorded by policy years. King accordingly assumed that the rate of mortality for the calendar year of entry, which applied on the average to only a fraction of the first policy year, would continue for the whole of that year and that the rate for the tth calendar year after entry would apply to duration t. He pointed out that this resulted in an understatement of the rates.

Subsequently, in 1879, T. B. Sprague constructed a complete set of select tables based on the H^M data. He assumed that the experience of the calendar year of entry related to the first half of duration 0, that of the calendar year after entry to the period between exact durations $\frac{1}{2}$ and $1\frac{1}{2}$, and so on. He then obtained the rates for integral durations by interpolation. The $H^{M(5)}$ table was used for the ultimate rates and the select rates for the first five years after entry were run into this table, notwithstanding the fact that it excluded only part of the data applicable to duration 4. Not only did Sprague introduce the notation now generally used for select tables, but his paper was an important contribution to the theory of the subject and resulted in the final acceptance by actuaries of the principle of select tables.

8. British Offices' Life Tables (1863-93).

In 1893, the preliminary work was begun on the most comprehensive investigation which had been attempted up to that time. The investigation is notable for the care with which it was carried out and the great attention which was paid to detail. Much interesting and valuable information is given in the official account of the investigation published in 1903 (*An Account of the Principles and Methods...*).

(i) *Treatment of data.*

Full details of the scope of the experience will be found in Appendix A of the official account in which the instructions to the contributing offices were given. Cards were returned for all policies which had been in force at any time during the period 1863–93, but data for the years before 1863 were excluded in order to secure a higher degree of homogeneity. According to modern standards, thirty years is of course too long a period to achieve this object. The dates of birth, entry and exit were recorded on the cards and comprehensive rules were made to ensure uniformity of practice between one office and another in dealing with cases where the correct data were doubtful. Lives subject to extra risks at the date of entry were excluded. The cards were marked to show the class of the policy and whether with or without profits. Different coloured

cards were used for male and female lives and different forms of cards for lives who entered before and after 1st January 1863.

The policy year method was adopted and lives were traced from their policy anniversaries in 1863 to those in 1893, the age at entry being taken as the nearest age at that date. An account of the investigation made to determine the most satisfactory method of recording the age at entry is given on pp. 37–8 in Appendix C and is well worth studying. For beginners and enders, exact durations were recorded on the cards by the operators; for deaths the curtate durations were entered. "Withdrawals" were kept separate from "terminations", the former cause of exit including surrenders and lapses, while the latter included maturities of endowment assurances and other causes not due to the action of the policy-holder. The object of making this distinction was to obtain data for calculating lapse rates.

A very thorough investigation was made in order to ascertain how the fractional exposures in the policy year of withdrawal could best be estimated. Eventually, after examining a sample of the data, the committee devised a method—described as a *modified nearest duration method*—which took into account the fact that a large proportion of the withdrawals occurred at the end of the thirty days of grace allowed for payment of premiums. It is unlikely that so thorough an investigation into this question will ever be made again and, if for no other reason, the account given on pp. 39–46 and in Appendix M deserves attention.

(ii) *Results of the investigation.*

A comparison of the mortality rates for the different classes of policy revealed significant variations. Of the principal classes, endowment assurances with profits showed the lowest rates, then whole-life with profits subject to a limited number of premiums, whole-life with profits, whole-life without profits and temporary assurances in that order; that is to say, other things being equal, the lower the rate of premium the higher was the rate of mortality. This feature was largely attributed to the tendency for a healthy life to back his chance of survival by effecting a policy under which he would not be too severely penalized if he enjoyed a long

life. An endowment assurance, for instance, is not an attractive contract to a man who does not expect to survive to the maturity date, for on previous death the benefits would in general be no greater than would be secured under a whole-life policy for which a lower rate of premium would be payable. The same argument applies, though with less force, in assessing the merits of an ordinary whole-life policy and a policy under which the number of premiums is limited. Again, a with profit policy has a greater appeal to the healthy than to the damaged life because the longer he survives, the greater will be the benefits.

A further reason suggested for the heavy mortality in the whole-life non-profit and temporary assurance classes was the fact that a policy of one or other of these types is usually effected when collateral security is required for a loan or in connexion with the purchase of a reversion or life interest, and that lives requiring policies in these circumstances are not so healthy as lives effecting policies for provident reasons or as investments.

There was, however, another important factor operating which largely vitiated the comparison between whole-life and endowment assurance mortality. During the years 1863–93 endowment assurances were growing in popularity and there was a marked increase in the proportion of lives effecting these policies. As a result, the endowment assurance business related on the average to a later period than the whole-life business and this factor, combined with the improvement in mortality rates with the passage of time, might easily have been responsible for the greater part of the difference in mortality revealed by the experience.

(iii) *Treatment of selection.*

Graduated tables were constructed for the main classes of business only, the principal select table being based on the whole-life with profits data. The effect of temporary initial selection (described simply as selection in the official account) was investigated by comparing values of $e_{[x-t]+t}$ for quinquennial age groups and for quinquennial values of t from 0 to 25. The expectation of life was used for this purpose to avoid the difficulty of comparing rates for individual ages or groups of ages due to the presence of random

errors. How far e_x is suitable for such comparisons will be discussed more fully in Chapter XVIII.

The comparison suggested that temporary initial selection had for all practical purposes spent its force by the end of ten years and it was decided to construct a table with a ten year select period. This is known as the $O^{[M]}$ table. The possibility that the length of the select period had been overestimated owing to the influence of spurious selection is not mentioned in the official account, although, as indicated in (XIII, 6), this view is now held by a number of actuaries. A very thorough examination of the data would be necessary if we were to discuss the problem further and this is hardly feasible within the limits of this book.

In the official account attention is drawn to one special feature in the table of $e_{[x-t]+t}$, i.e. the light mortality experienced by lives entering at ages 20–29. This feature was so pronounced that the expectation of life over a considerable range of attained ages was greater for entrants at ages 20–29 than for lives of the same attained ages who had entered more recently at older ages. No reason was given for this phenomenon and it may be instructive to put forward some possible reasons in the light of our study of selection, even although it is not practicable for us to investigate the correctness of the reasons suggested.

In the first place, it will be noted that the phenomenon under review is not consistent with the view that the secular improvement in mortality depends primarily on the calendar year of entry, for the lives contributing to $E_{[25]+20}$ entered earlier on the average than those contributing to $E_{[35]+10}$ (assuming an even distribution of the exposed to risk over the years 1863–93 the average years of entry would have been 1858 and 1868 respectively). The following are some of the influences which may have been responsible.

(*a*) The standard of selection exercised by the offices may have been higher at ages 20–29 than at older ages.

(*b*) At the younger ages, life assurance may have appealed more strongly to the class of life likely to experience light mortality than was the case for entrants at older ages.

(*c*) The proportion of entrants at young ages may have increased over the last half of the century, with the result that the average

calendar year of exposure for these entrants was later than for entrants at older ages. This feature, combined with improving mortality with the passage of time, would throw up lighter mortality for the young entrants.

Select tables are required primarily for calculating rates of premium. Strictly speaking, they should also be used for the periodic valuation of the policies on the books of an office, but the increase in work which would be involved if policies had to be grouped by duration as well as by class and age would be considerable. Moreover, for valuation purposes, we are less concerned with the correctness of the reserves set up for individual policies than with the adequacy of the total reserves. It is therefore customary to use an aggregate or an ultimate table for this purpose.

It was accordingly decided to construct from the whole-life with profit data an aggregate table (known as the O^{M}) and an aggregate table excluding the first five years' duration (known as the $O^{M(5)}$). These tables are comparable with the H^{M} and the $H^{M(5)}$ tables. They showed a marked improvement in mortality over the earlier experience, especially at the young ages. The slopes of the mortality curve are therefore steeper and the reserve values calculated therefrom are in general greater than for the H^{M} and $H^{M(5)}$ tables.

A select table, in which initial temporary selection was assumed to last for five years, was also constructed from the whole-life non-profit data. This table (known as the $O^{[NM]}$) was largely used for non-profit premiums, just as the $O^{[M]}$ was used for with profit premiums.

(iv) *Annuitants' experience.*

It is convenient to mention at this stage the investigation into the mortality of annuitants which was carried out at the same time as the investigation into the mortality of assured lives and which followed much the same lines. Select tables were constructed in which selection was traced for five years, although it was not considered that its effect had completely disappeared in that period. Males and females were dealt with separately and the resulting tables were called the $O^{[am]}$ and $O^{[af]}$ tables respectively. Aggregate tables were also prepared. Separation of the sexes is much more

important for annuitants than for assured lives, in view of the much larger proportion of females in the former class and the fact that at the ages with which we are chiefly concerned, i.e. from 50 upwards, female mortality is much lighter than male mortality. The number of females in recent experiences has been about three times the number of males.

Surrenders of annuities are very rare and in this investigation lives who withdrew were excluded altogether from the experience, except those who effected new contracts when the old ones were given up. This process is theoretically unsound, for the corresponding deaths could not be excluded since it was obviously impossible to say which lives would have withdrawn if they had not died previously. It is unlikely that this fact was overlooked and there was no doubt a good reason for the decision, but the official account is silent on this point. The number of years' exposure excluded was very small and the operators were presumably satisfied that the error was negligible.

(v) *General comments.*

In the light of more recent knowledge, the 1863–93 tables have been widely criticized. It has been argued that the inclusion of data from a period as long as thirty years introduced a considerable degree of heterogeneity and probably an element of spurious selection. This is undoubtedly true and it is unlikely that any future investigations into life office mortality will deal with so long a period, particularly as the increased volume of life assurance business enables us to obtain extensive data from an investigation covering only a few years. It has also been suggested that the very thorough and painstaking investigations to some extent defeated their own object in that they delayed publication till ten years after the end of the period from which the data were drawn. On the other hand, these investigations threw light on a number of questions which had not previously been considered fully and provided valuable information for the benefit of future investigations.

Whatever criticisms may be made, the tables were undoubtedly a marked improvement on earlier tables and they have proved of great practical value to life offices.

9. The Continuous Investigation—A 1924-29 Table.

Another large-scale investigation was contemplated in 1914, but with the outbreak of war it had to be postponed. A start was not made until 1924 as it seemed likely that the effects of the war would persist for some years after its conclusion.

It was known that there had been a marked improvement in mortality since the 1863–93 investigation and, with the prospect of further improvement in the future, the publication of tables within a few years of the close of the investigation period was considered to be of primary importance. Simplicity was therefore the keynote of the new experience. Another factor which influenced the choice of method was the desirability of keeping a continuous record so that later investigations could be made without having to set up new machinery. With these objects in view, the census method was chosen, as this method not only involves comparatively little work for the contributing offices but is particularly convenient for a continuous investigation.

All students of mortality tables are recommended to read the official account of the investigation and of the meetings devoted to discussing it (*Mortality of Assured Lives* 1924–29). We shall mention here only a few salient features.

(i) *Treatment of the data.*

Separate returns were made for whole-life and endowment assurances and for with and without profit policies. In each of these four groups the data for medical and non-medical cases were shown on separate sheets, unless the proportion of the latter was less than 10 per cent of the whole and it was inconvenient for the office to separate the two classes. Lives accepted on special terms were excluded, as were female lives unless the proportion to the whole was less than 5 per cent.

Offices were allowed to choose their own methods of recording the ages of the lives included in their returns. This latitude was granted to enable them to use the same age-groups as in their valuation schedules. It caused however a good deal of extra work for the operators, who had of course to adjust the data to make them consistent with the method of age-grouping chosen for the

investigation, i.e. age nearest birthday on the appropriate date in the calendar year for the populations and age nearest birthday at death, a method by which the data for nearest age x give the rate of mortality for age $x-\frac{1}{2}$. Moreover, some errors discovered in the returns would have been avoided if a rigid method of age-grouping had been insisted upon and it is doubtful whether in the long run the latitude granted achieved any saving of time. There is the further point that the adjustments necessary to achieve consistency involved assumptions which entailed some loss of accuracy, although the errors introduced were almost certainly negligible.

The original intention was to construct tables from the data for the years 1924–26, but it was subsequently decided to await the experience of the years 1927–29, as it was considered that the mortality of the former three years was abnormally light. Actually, there proved to be little difference in the mortality between the two periods and in this connection it is worth noting that there was no major epidemic at any time during the six years. Further delay was caused by the discovery of serious errors on the part of some of the offices which necessitated a good deal of additional work and, as a result, tables were not published until 1933. Nevertheless, this was a great improvement on the time required for the 1863–93 investigation, i.e. four years from the closing date of the investigation instead of ten.

(ii) *Results of the investigation.*

Tables of ungraduated rates for the four main classes of policy were prepared for each of the periods 1924–26 and 1927–29, medical and non-medical business being dealt with separately. The tables showed that, while generally speaking the endowment assurance class had continued to enjoy lighter mortality than the whole-life class, there had been a change in the relative mortality of with and without profit business and the latter class now appeared to experience somewhat lighter mortality. The conclusion was therefore reached that class of policy does not exercise a stable influence on mortality and, as there was no reason to expect greater stability in future, it was felt that there was no point in constructing separate tables for the different classes. Several possible causes of the varia-

tions were advanced in the official account and in the discussions, but up to the present no convincing evidence in support of any of these explanations has been published. We can therefore go no further than the statement made in the official account: "...we cannot assert that the mortality of any particular class is more homogeneous than that of the whole business of an office or than that of the offices as a whole."

As regards medical and non-medical business, the data for the latter class were scanty and the differences revealed could not be regarded as significant.

In considering the effect of the factors referred to above, it should be remembered that the offices themselves exercise some influence. Proposers for whole-life assurances who are not acceptable at the ordinary rate of premium under that class may be diverted to the endowment assurance class by the offer of policies on normal terms. This practice tends to increase endowment assurance mortality, for, although the extra risk may be too small at the ages up to the maturity age to justify an extra premium, it may not be altogether negligible. Again, it is customary to demand medical examination when a proposer under the non-medical class is a doubtful risk and the cases subsequently accepted at ordinary rates in these circumstances may easily be below the average of the lives accepted after medical examination as a whole.

(iii) *Treatment of selection.*

No attempt was made to investigate fully the duration of temporary initial selection. Data subdivided according to duration were available for only five years after entry and the effects of selection had not apparently worn off in that period. The possibility of spurious selection is mentioned in the official account. The decision to trace selection for only three years is justified by the object of the tables, i.e. the calculation of office premiums and reserve values, for the effect of slightly underestimating the duration of selection could hardly have been of any practical importance from this standpoint, except possibly for temporary assurance premiums.

(iv) *General comments.*

The continuous investigation has been criticized mainly on the ground that it has done little to solve any of the general problems of mortality in which actuaries are interested. The duration of initial selection, variations in mortality between different classes of assurance or between medical and non-medical business are examples of the questions which remain unanswered. On the other hand, the investigation has at least performed the valuable though negative service of showing that conclusions which had been reached as the result of previous investigations—notably that of 1863–93—are not necessarily true under modern conditions or at least cannot be supported by the available statistical evidence.

What is much more important, the investigation has focused attention on the nature of the difficulties still to be overcome—which is a valuable step in the direction of finding solutions. In particular, emphasis has been laid on the importance of statistical errors due to heterogeneous data. This problem was neglected by the investigators of the 1863–93 experience, who concentrated on the detailed methods of calculating E_x and the errors involved in the various assumptions made about ages, durations, etc.

One cause of heterogeneity to which attention was drawn was the wide difference between the experiences of the contributing offices, arising presumably from their different standards of selection and the different districts and social classes from which their policy-holders were drawn. This led to the publication of tables based on the data contributed by offices experiencing light and heavy mortality respectively and has added considerably to the value of the investigation. (For a full discussion of this question, see the paper by Elderton, Oakley and Smither, *J.I.A.* Vol. LXVIII.) The variations between offices are undoubtedly a major cause of heterogeneity. This suggests that future progress may be made by grouping the data of offices which have similar standards of selection and draw their policy-holders from the same social grades. It would not, however, be a simple matter to choose suitable groups, for when the experience of an individual office is under review random errors may have a disturbing effect.

The accumulation of more and more data as the years pass will

be of material assistance to the investigators, although time is itself a cause of heterogeneity, both because of general changes in mortality and changes in the practice of particular offices. In the meantime, it must be admitted that our knowledge of these and similar problems of mortality is still in its infancy and much research remains to be done.

The investigation has been further criticized on the ground that, by using the census method and grouped returns in schedule form instead of obtaining particulars of individual policies, the scope of the enquiry was unduly restricted and many possible lines of research excluded *a priori*. For example, it was suspected that one cause of variation in the mortality of offices was the varying proportions of "old" business on their books—an office with a large proportion of business effected many years ago being expected to show a heavier mortality than an office most of whose business had been written recently. The returns gave no information which would assist in answering this question. Many other lines of investigation, such as the effects on mortality of occupation, physical impairments, sum assured, etc., were also automatically excluded. In reply to this criticism, it can be urged that the main purpose of the investigation was to produce as quickly as possible an up-to-date standard table of the mortality of assured lives and that the A 1924–29 table is fulfilling this function satisfactorily. A study of the discussions on the investigation will show how opinions differ as to the relative importance of these different aspects.

CHAPTER XVI

FORECASTING RATES OF MORTALITY. GENERAL THEORY OF PROJECTION

1. General definition of the rate of mortality.

Hitherto we have defined the rate of mortality as the ratio of the number of deaths occurring among a body of lives to the corresponding exposed to risk. This definition is satisfactory when applied to the crude rate of mortality (i.e. the unadjusted rate calculated from the crude data), but is not suitable as it stands when it refers to the graduated rate. It is desirable to have a definition which is generally applicable and the following new definition will accordingly be adopted.

The rate of mortality at age x among a group of lives, who are not subject to any causes of increment nor to any causes of decrement except death, is the probability that a life aged x will die before attaining age $x+1$.

This definition applies equally to the crude and to the graduated rates. Suppose, for example, that we refer to the rate of mortality at age x experienced in the years 1924–29 by the whole-life with profit policyholders contributing to the Life Offices investigation. The definition states that we mean the probability of any one of these lives aged x chosen at random being one of the recorded deaths between ages x and $x+1$, which is consistent with the original definition. On the other hand, if the mortality according to the A 1924–29 table is mentioned, the application of the definition is self-evident and the columns of l_x and d_x give us what we require without reference to actual deaths or numbers exposed to risk.

2. Forecasting rates of mortality.

The new definition is capable of even wider application. In the last few chapters we have seen how often we wish to use the experience of the past as a guide to the future. In so doing we admit the possibility of forecasting future rates of mortality.

Whenever we use mortality rates to measure the probability of future events (as, for example, in calculating rates of premium), we are committing ourselves to an estimate of future mortality. If we employ rates based on past experience without modification, as we frequently do, we are in effect assuming that the past rates will be experienced in the future. This assumption may well be without justification.

We have already referred to the improvement in mortality which has occurred during the last hundred years as a result of the advances made in preventing and combating disease. Much remains to be done in this direction and there is no reason why, in spite of occasional temporary setbacks, the general level of mortality rates should not continue to fall for many years to come. In these conditions there are obvious objections to the calculation of probabilities of future events from a mortality table based solely on past experience, even if the experience is recent. There are many reasons for believing that a better estimate of future mortality could be made by reference to the *trend* of values calculated in the past. The hypothetical values so obtained would not, of course, be based on any particular set of deaths and our earlier definition of the rate of mortality is therefore unsuitable. They do, however, signify the probabilities of dying which we estimate will apply to a particular class of lives at a particular time; the new definition is therefore applicable without modification.

3. Generation mortality.

At a very early stage in this book, it was pointed out that, for practical reasons, we cannot trace from birth to death a group of lives born in the same calendar year. Let us examine in more detail the difference between this process and our substituted expedient of observing a group of lives of all ages over a short period of time.

If in the year 1940 we wished to calculate an annuity value, a rate of premium or a reserve value for a life aged 30, the rates of mortality required would be those which lives of the appropriate class would experience at age 30 in 1940, age 31 in 1941, age 32 in 1942 and so on—in other words the mortality likely to be experienced by the

generation born in 1910. Similarly, for a life aged 35 in 1940, we should use the generation mortality applicable to lives born in 1905. To make proper allowance for the effect of time on mortality, we therefore require a separate mortality table for each attained age in 1940, i.e. a separate table for each generation.

In contrast to this, a table based on the mortality of a single calendar year is a cross-section of the generation tables, containing as it does a single rate from each of them. In the same way, a mortality table taken from the experience of five calendar years is based on rates of mortality each of which is the mean of the rates applicable to five generations. Thus, in an experience covering the years 1935–39, for example, the rate of mortality at age 30 would be the mean value of q_{30} for the generations born in the years 1905–09, weighted by the exposed to risk drawn from the different generations, while the rate at age 35 would refer to the years of birth 1900–04. It is evident, therefore, that a mortality table of this type does not demonstrate the mortality throughout life of any particular body of lives and the successive values of l_x do not show the number of survivors of a particular generation from age to age. Its limitations as a measure of future mortality are also apparent. Suppose, for example, that the experience of 1935–39 referred to above were available to calculate the future mortality of a life aged 30 in 1940. The table gives us a value of q_{30} relating to years of birth 1905–09, or say 1907, a value of q_{35} relating to years of birth 1900–04, or say 1902, and so on, whereas in each case we really require the rates for year of birth 1910. No matter how recent a mortality experience is available, therefore, the estimates that it provides of the future become increasingly divergent from what is required.

4. Tables of generation mortality.

The preceding paragraph naturally leads us to ask if there is any method of constructing a mortality table which will give us a better measure of generation mortality.

Suppose that, for a particular class of lives, there are available the rates of mortality experienced at quinquennial intervals for

many years past. We can then draw up a table of the rates for quinquennial age groups in the following form where q_x^N = rate of mortality at age x for lives born in calendar year N.

Central Attained Age	Calendar Year of Experience							
	1840	1845	1850	...	1935	1940	1945	1950
20	q_{20}^{1820}	q_{20}^{1825}	q_{20}^{1830}	...	q_{20}^{1915}	q_{20}^{1920}	q_{20}^{1925}	
25	q_{25}^{1815}	q_{25}^{1820}	q_{25}^{1825}	...	q_{25}^{1910}	q_{25}^{1915}	q_{25}^{1920}	
.	.	.	.	...	.	.	.	
.	.	.	.	...	.	.	.	
.	.	.	.	...	.	.	.	
.	.	.	.	...	.	.	.	
.	.	.	.	...	.	.	.	
.	.	.	.	...	.	.	.	
.	.	.	.	...	.	.	.	
95	q_{95}^{1745}	q_{95}^{1750}	q_{95}^{1755}	...	q_{95}^{1840}	q_{95}^{1845}	q_{95}^{1850}	
100	q_{100}^{1740}	q_{100}^{1745}	q_{100}^{1750}	...	q_{100}^{1835}	q_{100}^{1840}	q_{100}^{1845}	

Each vertical column gives the rates applicable to a particular calendar year passed through, while the rates experienced by a particular generation are given by the appropriate diagonal reading downwards and from left to right, e.g. the figures for year of birth 1820 are given in the diagonal beginning at the top left-hand corner of the table. Each horizontal line shows the trend of the rates for a particular age group with the passage of time.

What we require is an extension of this table to the right, so that we may have a measure of the mortality to be experienced in future. Provided that some fairly definite trend is discernible in the past rates, there is no reason why such an extension should not be attempted either by the use of mathematical formulae or by a more rough and ready method. Clearly the resulting table will be hypothetical, since it will not be based on any actual experience and will involve forecasts of future mortality. It does not follow that it may not be a more suitable instrument for certain purposes than a table based exclusively on rates which have been actually experienced in the past. The process of estimating future rates of mortality by reference to those experienced in the past is known as *projection* or *forecasting*.

5. General theory of projection.

An examination of the above table suggests that there are two ways of tackling the problem of projection, i.e. along either the horizontal or the diagonal lines. These two methods involve rather different assumptions about the trend of rates of mortality. The first assumes that the rate of mortality at a given age will vary with the time of observation in some recognizable way: the second assumes that the values of q_x^N for a given value of N will follow some definite law or will at least vary in such a way that, given the values of q_x^N, $q_{x+t}^N \ldots q_{x+nt}^N$, we shall be able to predict the unobserved values $q_{x+\overline{n+1}t}^N$, $q_{x+\overline{n+2}t}^N$, etc. In neither method, however, would the projection be made along an individual line without any regard to the adjacent lines. Whichever method were adopted, therefore, both the horizontal and the diagonal trends would in effect be taken into account. The difference would be largely one of emphasis, depending on which trend was considered first.

Clearly the major difficulty in practice will be to obtain adequate or satisfactory data. The very extensive table of para. 4 requires observations at quinquennial intervals for a whole century and even then we have relatively few values on which to base our projection for certain ages and generations. When we come to consider the methods adopted in the two investigations where projection was first employed, viz. the Office Annuitants' Experience 1900–20 and the Government Annuitants' Experience 1900–20, we shall see how this difficulty can be avoided. At this stage we shall consider how projection might be handled if adequate data were forthcoming—a proviso which we hope will eventually be satisfied through the medium of the Continuous Mortality Investigation.

The theoretical aspect of the subject has been dealt with in a number of papers of which the most important are those by Davidson and Reid (*T.F.A.* Vol. XI, p. 183) and Derrick (*J.I.A.* Vol. LVIII, p. 117) which were read on the same night. Both papers take the view that the most satisfactory way of obtaining a smooth progression of values is by projecting the rates for individual generations (or groups of generations) rather than individual ages (or groups of ages), i.e. along the diagonal instead of the horizontal lines in the table. They approach the subject, however,

in rather different ways. Davidson and Reid use a method of curve fitting based on a mathematical formula for q_x; Derrick employs a statistical argument using figures drawn from the National Life Tables. Both papers deserve careful study.

6. Projection by curve fitting.

On the general theory of generation mortality, the arguments put forward by Davidson and Reid in the paper referred to above (and in the somewhat simpler account given by Reid in *A.S.M.* 3) may be briefly summarized thus:

(1) Our knowledge of the rate of mortality at a particular age for a given class of lives in a given year is drawn from (*a*) the rates of mortality at that age and at adjacent ages experienced in the year in question and (*b*) the results of earlier experiences based on lives of the same class. Mortality experiences of the traditional type take account of (*a*) to arrive at the graduated rates, but little or no weight is given to (*b*). A better measure will be obtained if we use all the available data.

(2) If a law of mortality exists, it is more likely to relate to rates of mortality based on consecutive human life than to rates obtained from observations over a short period of time.

(3) Such a law would itself be subject to variation from one generation to another, reflecting the varying factors influencing human survival.

(4) Though these individual factors are varying and temporary in character, the resultant effect will not be entirely haphazard. It is therefore permissible to hope that the variations will follow a definite trend and be capable of representation by a mathematical formula or by a smooth curve.

For illustration the authors used the Makeham formula $\mu_x = A + Bc^x$, where A, B and c are constant for each generation, but vary from one generation to another. By a suitable choice of values for A, B and c, a curve of this form would be fitted to the available rates for each generation. For the earlier generations, assuming data to be available for many years past, there would be a series of rates covering a wide range of ages; the data for the later

values of the constants would be unreliable. Thus, for an investigation in 1940, the generation born in 1860 would supply rates for ages up to 80, whereas the generation born in 1910 would only have reached age 30 in 1940.

The next stage of the method involves a comparison of the values of the constants for the different generations. If, as we expect, there is a definite trend discernible for each constant, the values could be submitted to a process of graduation to remove random errors—a process which is only justified if there exists a smooth underlying curve which would apply if unlimited data were available.

It will be seen that, although the projection would be made along the diagonal lines in the table of para. 4, the trend of the values read horizontally would in effect be taken into account by the process of graduation, which involves the linking up of adjacent values of the constants.

7. Statistical approach to the theory of projection.

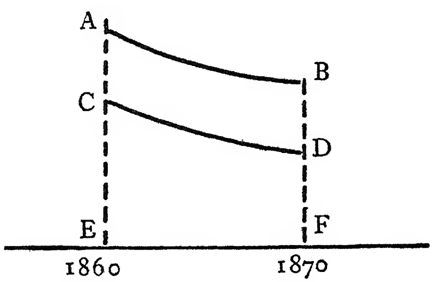
Although Davidson and Reid illustrated their theory of generation mortality by numerical applications, they were careful to explain that these were mere illustrations and were not intended as statistical evidence to support it. Derrick, on the other hand, approached the problem of projection from the purely statistical point of view. The necessary figures were obtained from mortality tables based on the national statistics of England and Wales, which extended as far back as 1841 and accordingly provided sufficient evidence to permit of an investigation into the trend of general population mortality. It must not be concluded that the same results would have been obtained if the data had been confined to assured lives or annuitants. The mortality rates corresponding to nine calendar years between 1846 and 1923 were first plotted graphically for a number of age groups, the abscissae representing calendar year of experience and the ordinates rates of mortality. This produced a series of rough curves representing the trend of q_x for various values of x. The curves did not indicate any general tendency in common to them all.

The next stage was to make the abscissae represent year of birth instead of year of experience and to plot the values of q_x again.

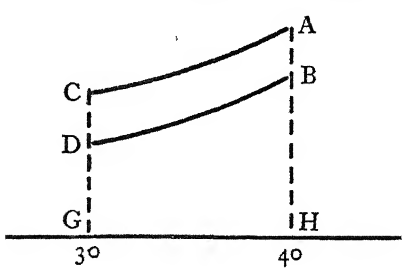
The effect of this was that values of q_x applicable to a given calendar year, which had previously lain on a vertical line, were now extended roughly in a diagonal line descending from left to right with decreasing age. In other words, the curves of q_x were moved horizontally to the right, except for that corresponding to the oldest age group and the younger the age group, the greater was the distance moved.

It was then found that the curves showed a certain degree of parallelism, which suggested that generation influences were an important factor in mortality improvement. The correctness of this conclusion may not be at once obvious; it may be more easily understood if we appreciate that, if the curves of q_x applicable to different values of x are parallel when arranged as above, the generation curves will also be parallel.

To illustrate this, suppose that the following diagram represents a small section of the two curves q_{30} and q_{40} for years of birth between 1860 and 1870.



The point A represents q_{40}^{1860}, D represents q_{30}^{1870} and so on. Now, if we were to plot the appropriate generation curves, the corresponding sections would run roughly as follows:



The ordinates of the points A, B, C, D are the same in both diagrams, but CA and DB are now part of the curves applicable to the generations born in 1860 and 1870 respectively. When the curves AB and CD are parallel, as in the first diagram, $AC=BD$ and $AE-CE=BF-DF$. But as the ordinates of the points are the same in both diagrams, $AH=AE$, $BH=BF$ and so on. Hence $AH-CG=BH-DG$, so that $AH-BH=CG-DG$ and $AB=CD$. The curves in the second diagram are therefore parallel over the section under review.

It must not be supposed that the figures put forward by Derrick provide a convincing proof of the theory of generation mortality. The author himself went no further than to claim that the principle of projecting generation rates should in future receive consideration equal to that accorded to the method of projecting the rates for individual ages without regard to the mortality experienced at other ages. Nor was it suggested that mortality improvement depends solely on year of birth and is independent of all other factors. The problem is rather that of deciding which of the many factors, other than age, influencing mortality rates is the most important; the view expressed both by Derrick and by Davidson and Reid is that the calendar year of birth probably exercises the greatest influence. Whichever view be taken, it is important to realize that the two methods of projection, i.e. horizontally and diagonally, do not differ in principle. For practical purposes, the question concerns us in so far as it indicates the line along which the long term trend of the rates will be most easily discernible.

8. General observations.

Although it is hardly within the province of the actuary to advance arguments, other than those based on statistics, for or against the theory of generation mortality, it is interesting to consider some of the arguments which have been put forward. These are largely based on conjecture and are mentioned here merely to show that the problem raises very wide and controversial questions and is not of interest to actuaries alone.

In favour of the theory, it has been suggested that the early years of a child's life are of great importance in determining his or her

future well being; hence conditions relating to such matters as pre-natal care, diet and hygiene, existing in or about the year of birth, are of primary importance, not only in their effect on mortality in infancy and childhood, but over the whole subsequent life-time of the generation. Improvements in these conditions would not, on this assumption, affect the general trend of the generation curve to a marked extent, but would lower the level of the curve compared with earlier generations. On the other hand, such changes would affect the curves applicable to individual ages in different calendar years and this would tend to obscure the relationship between these curves.

In opposition to the theory, it might be argued that some improvements in medical science bring about an immediate reduction in the rate of mortality at all ages, e.g. the discovery of a new drug. The age curves would all show the effect of this factor at one point, i.e. in the neighbourhood of the calendar year when it first began to operate, but the generation curves would be affected at different ages and would differ in shape, age by age, until the youngest generation existing at the date of the change had expired. In that event, it might be more satisfactory to project along the age curves than along the generation curves. The problem is a fascinating one but this aspect of it is outside the scope of this book.

In conclusion, we should remind the student that there may be influences operating on mortality other than age or generation and that in certain circumstances these may be of greater weight. For example, in an assurance or annuity experience a change in the type of entrant might profoundly affect mortality. The rates of mortality might in that case vary with year of entry rather than year of birth. If it were thought likely that this feature would persist in the future, projection might be carried out on a year of entry basis. In the following chapter we shall discuss this problem in the light of a recent investigation of Life Office annuitants.

CHAPTER XVII

MORTALITY TABLES OF ANNUITANTS

1. The idea of forecasting has not found universal favour with actuaries, some of whom feel that it is better to employ rates which have actually been experienced in the past, rather than to embark on forecasts which, in our present state of knowledge, must involve a large element of guesswork. In particular, they argue that it is out of place to introduce the element of forecasting into a standard table and that such a table should be based on ascertained facts, leaving the individual to make such adjustments as seem desirable. So far as the mortality of assured lives is concerned, this view has been encouraged by the steady improvement in mortality during recent years. As a consequence of this improvement, the use of past rates of mortality without modification has provided a useful margin to cover the risk of losses due to temporary increases of mortality rates or even to unexpected reductions in rates of interest.

For annuitant mortality, on the other hand, it is obviously necessary to provide a margin against future improvement and the principle of forecasting in annuitant tables has accordingly been fairly generally accepted. The only two major investigations in this country where projection has been employed are the Office Annuitants' Experience 1900–20 and the Government Annuitants' Experience 1900–20. These we shall now discuss with particular reference to the methods of projection employed.

2. Collection and tabulation of data.

(*a*) *Office Annuitants' Experience* 1900–20. Returns were made by the offices on cards which were in a suitable form for subsequent use as machine cards. It was decided to trace selection for not more than five years and to divide the twenty years of the investigation into three periods in order to ascertain whether any significant change had taken place in mortality rates over the whole period. Instead of recording merely the dates of birth, entry and exit and the mode of exit, the offices also inserted the age last

birthday at entry and the duration (or, when the duration was 5 or more, the age) at which the life came under and passed out of observation in each of the three periods.

The experience extended from the anniversary of entry in 1900 or from subsequent entry to the anniversary in 1920. The dividing points for the three periods were the anniversaries in 1907 and 1914. The exact durations for beginners and enders and the curtate durations for deaths and withdrawals were inserted, the corresponding ages being obtained by adding the appropriate durations to the ages last birthday at entry. Withdrawals were assumed to be evenly spread over the policy year. The exposed to risk formulae can easily be established; although the actual process differs slightly from that given in Chapter V, the results are the same.

The tables based on the experience for males and females are known as the $a(m)$ and $a(f)$ tables.

(*b*) *Government Annuitants' Experience* 1900–20. The observations covered all lives on which Government life annuities had been granted. The experience under Savings Bank annuities had not previously been investigated and the results for these lives and for other Government annuities were set out separately. It was found that the mortality of the two classes was sufficiently similar for them to be combined in the final table. As in the Office Annuitants' Experience, the lives were traced from the anniversary of the date of purchase in 1900 or from subsequent entry to the anniversary in 1920 and the data were subdivided into three periods. Male and female lives were treated separately and selection was traced for a period of 5 years after purchase.

3. Variation of q_x with year of birth.

In both investigations it was clear from a preliminary inspection of the figures that there had been a marked improvement in mortality since the publication of the previous comparable table. Rates based solely on the 1900–20 data would therefore be unsafe as a measure of the selling price of annuities, as it was probable that mortality would continue to improve. In each experience, therefore, it was decided to regard q_x as a function, not only of x, but of N, the calendar year in which age x was reached, and to

investigate the form that q_x^N would take considered as a function of N. The following points had to be taken into account.

(*a*) The only past values of q_x^N which could legitimately be used were those obtained from past investigations of similar annuitant lives. In each case, therefore, the only available values of q_x^N were those based on the new experience and on one former investigation, i.e. for life office annuitants, the 1863–93 experience and for Government annuitants, an earlier experience (1875–1904).

(*b*) The number of constants in the expression for q_x^N could not be greater than the number of available values of q_x^N; otherwise there would not have been sufficient equations to solve for the values of the constants.

(*c*) The graph of q_x^N could not be a straight line or negative values would be obtained for sufficiently large values of N. In fact, q_x^N must tend to a definite limit as N tends to infinity.

Clearly therefore, the choice of a function for q_x^N was limited and a geometrical formula of the type $q_x^N = \alpha_x + \beta_x \gamma_x^N$ was a fairly obvious selection. In this expression, α_x, β_x and γ_x are functions of x only and for a given value of x remain constant as N varies. We must have $\gamma_x < 1$ or q_x^N will not decrease as N increases.

The idea was to ascertain for each value of x the appropriate values of α_x, β_x and γ_x which would satisfy the above relation in the calendar years for which values of q_x^N were available. The value of q_x^N for any future calendar year would then be obtained by inserting the appropriate value of N in the expression $\alpha_x + \beta_x \gamma_x^N$. In effect, therefore, projection would be effected along horizontal and not diagonal lines.

The values of q_x^N from the two life office experiences were assumed to apply to the years 1880 and 1910 on the average and these were the only values based on past experience which were employed to determine the constants. A third value was obtained by fixing the limiting value of q_x^N and thus assigning a value to the constant α_x, since q_x^N tends to α_x when N tends to infinity. It was thought likely that α_x would lie between 50 and 75 per cent of the ultimate rate of mortality according to the 1900–20 experience; tests showed that variations in the percentage between these limits

would have little effect on the resulting annuity values. The value of α_x was eventually taken to be 63 per cent of the ultimate rate of mortality from the 1900–20 experience. It was then possible to calculate values of β_x and γ_x for all values of x and hence values of q_x^N for all values of N. Forecast rates on this basis were published for quinquennial age groups applicable to the calendar years 1925, 1935 and 1945. Complete mortality tables were not, however, constructed on this basis.

In the case of the Government annuity experience, three past values of q_x^N were used to obtain the constants for forecasting rates of mortality. One of these was taken from the 1875–1904 experience and was assumed to apply to 1890. The other two were obtained from the 1900–20 experience. For other purposes the data for the latter had been divided into three parts covering the periods 1900–07, 1907–14 and 1914–20 respectively, and the rates from the first and last of these periods were used for purposes of projection. They were assumed to apply to 1904 and 1918. It was found that the resulting values of α_x, the limiting rate of mortality, were usually about 80 per cent of the ultimate rates for 1900–20, a result which supported the use of a constant percentage independent of age in fixing the limiting rates for the Offices' table.

Theoretically, the next step in each investigation would have been to construct a complete table of forecast values for all ages and all calendar years from 1921 onwards. This table would, however, have been cumbersome and inconvenient for the calculation of annuity values; in both investigations, therefore, an attempt was made to produce by some practical device an instrument which would be more convenient in practice, but would nevertheless make a suitable allowance for the expected improvement in mortality. The solutions reached were on different lines and we shall discuss them separately and in detail.

4. Office Annuitants' Experience 1900–20: method of projection.

Probably the clearest account of the experiments which led to the final solution adopted in this table is given in Sir W. P. Elderton's paper *Forecasting Mortality*, read to the Scandinavian actuarial societies. It appears that, having approached the problem of fore-

casting on the lines already indicated, the actuaries carrying out the investigation had the idea of constructing a single mortality table based on the average rates of mortality which they estimated would be experienced by lives purchasing annuities in 1925. It was, of course, realized that such a table would soon be out of date, but it was obviously an advantage to be able to use a single table for all calculations of annuity rates made in a particular year, even if the single table had to be replaced or modified every few years.

The actual method of obtaining the average rates was simplified by an ingenious device. The 1863–93 data were used to determine separately for each value of x the average duration in force at the time of attaining age x. For example, the female annuitants attaining age 72 in the years 1863–93 entered on the average 12 years before reaching that age. It was assumed that these average durations would also apply to entrants in 1925 and that the latter entrants would accordingly reach age 72 in 1937 on the average. The value of q_{72} in the table intended for entrants in 1925 was therefore taken to be the forecast rate for age 72 in 1937. Similarly the female annuitants attaining age 62 in 1863–93 entered 10 years earlier on the average. The value of q_{62} for 1925 entrants was accordingly taken to be the forecast rate for age 62 in 1935.

The formula used to obtain the values of q_x required for the table was as follows, n_x being the average period between the year of entry and the year of attaining age x, according to the 1863–93 experience, q_x^N the rate experienced in calendar year N and q_x^∞ the limiting rate of mortality determined as explained in para. 3.

$$q_x = q_x^\infty + (q_x^{1910} - q_x^\infty) \times \left(\frac{q_x^{1910} - q_x^\infty}{q_x^{1880} - q_x^\infty} \right)^{\frac{15+n_x}{30}}. \qquad \ldots\ldots(1)$$

The reader should have no difficulty in deducing this formula by solving the appropriate equations for β_x and γ_x.

Formula (1) may be expressed in a slightly different form. q_x^{1880} and q_x^{1910} may be regarded as approximations to the average rates of mortality at age x applicable to entrants at all ages in $1880 - n_x$ and $1910 - n_x$ respectively. This can be expressed in symbols by substituting $q_x^{(1880-n_x)}$ and $q_x^{(1910-n_x}$ for q_x^{1880} and q_x^{1910}, the symbol $q_x^{(k)}$ denoting the rate of mortality applicable to entrants in year k.

This alternative form may convey the impression that the forecast is based on rates of mortality applicable to calendar years of entry. Clearly, however, the change in form does not in any way alter the nature of the formula which, as we have seen, is based on the rates experienced in 1863–93 and in 1900–20 and not on rates applicable to particular years or groups of years of entry.

The position would be different if it were assumed that the average period between entry and the attainment of age x had not been constant. Taking this period as n_x for the 1863–93 experience and n'_x for the 1900–20 experience, the years of entry corresponding to the years of experience 1880 and 1910 would be $1880-n_x$ and $1910-n'_x$ and the formula would become

$$q_x^{(1925)} = q_x^{\infty} + (q_x^{(1910-n'_x)} - q_x^{\infty}) \times \left(\frac{q_x^{(1910-n'_x)} - q_x^{\infty}}{q_x^{(1880-n_x)} - q_x^{\infty}}\right)^{\frac{15+n'_x}{30+n_x-n'_x}}.$$

This formula is a forecast based on approximations to the rates applicable to years of entry.

5. Office Annuitants' Experience 1900–20: suggested explanation of special features.

The actuaries next considered under what conditions this method would be theoretically correct, instead of being merely a convenient approximation. The necessary conditions were that the rate of mortality at age x should depend on the calendar year in which the annuity was purchased, irrespective of the calendar year in which age x was attained. This means that two groups of lives entering in 1925 at ages 40 and 80 would experience the same rate of mortality at age 90, although one group would reach this age in 1935 and the other not until 1975.

At first sight, it seems unlikely that these conditions would hold, but the existence of a peculiar feature in the data which could be explained thereby pointed to the desirability of further investigation. The feature in question was the failure of the select rates to run smoothly into the ultimate. A number of possible explanations may be advanced.

(*a*) Initial temporary selection might have lasted for more than five years, the period for which select data were available. As, how-

ever, the difference between $q_{[x-t]+t}$ and $q_x^{(t+1)}$ showed no tendency to disappear when t was increased from 1 to 4, this explanation would have required a considerable extension of the select period.

(*b*) Spurious selection might have been thrown up by the combination of improving mortality rates depending on year of birth—or, what is the same thing in effect, year of exposure—and variations in the distribution of the exposed to risk according to calendar year at successive durations and the same attained age. This type of spurious selection was discussed in (XIII, 5). In order that it may be present to a marked degree, there must be considerable variations in the distribution of the data. As these variations did not in fact occur this possibility can be ruled out.

(*c*) Spurious selection would be present irrespective of the distribution of the data, if the rate of mortality at age x were decreasing in such a way that q_x depended on the calendar year of entry and not directly on the calendar year in which age x was attained (XIII, 7). Under these conditions, spurious selection would be eliminated if the investigation were confined to the lives entering over a comparatively short period. Rates were accordingly calculated for the years 1900–20 in respect of the lives contributing to the select data in 1900–07 and it was found that, while the select rates did not run into the ultimate after five years, the differences were smaller than before. There was therefore, some corroborative evidence for alternative (*c*).

The authors of the report took the view that there was insufficient evidence of the continuance of temporary initial selection for more than one year after entry to justify the construction of select tables extending beyond the first year. They were influenced by the desirability of having an instrument which would be convenient to use and from this standpoint a strong case has been made for the process adopted. Admittedly, it has been suggested that the use of a one-year instead of a five-year select period might involve understatement of annuity values at the old ages by as much as 6 per cent, but in general the use of a one-year select period as a practical expedient has met with approval.

A good deal of controversy has, however, been aroused by the

suggestion advanced in the official report that the apparent selection revealed by the data was almost entirely spurious and was due to the dependence of the rates of mortality on calendar year of entry. It was impossible to produce convincing statistical evidence which would prove or disprove this view and, in the circumstances, it was inevitable that there should be differences of opinion.

If adequate data were available, it would be desirable to carry out a full investigation on the lines of the special investigation referred to in (*c*) above, i.e. an examination over a long period of the mortality of lives entering within the space of a few years. If the explanation given in (*c*) were correct, we should expect that after the first year or two q_x would be independent of the duration in force and the year passed through, but would differ from the values of q_x applicable to other calendar years of entry. Even if this were established, however, there are a number of different ways of accounting for the phenomenon.

The suggestion put forward by the authors of the official report is that it might be caused by changes in the class of person effecting annuities, either as a result of economic factors or for other reasons. They point out that if, during the years preceding the 1900–20 investigation, there had been an increasing tendency for annuitants to be drawn from classes experiencing comparatively light mortality, the result would be to throw up spurious selection in an experience in which the data were aggregated irrespective of calendar year of entry, but that this would not occur if the data for only a few calendar years of entry were segregated.

An alternative explanation is that, contrary to the view taken by the authors of the report, initial temporary selection might last more than five years after entry. The combined effect of this selection and of secular improvement in mortality rates (i.e. improvement associated with calendar year passed through) might be such as to make q_x apparently independent of duration in force and year passed through for lives entering over a short space of time. This possibility arises from the fact that secular improvement in mortality acting alone would result in reversed initial selection being exhibited by a group of lives entering in a given calendar year, since the data at duration $t+1$ would relate to a later calendar

year of experience than the data at duration t and would therefore be subject to lighter mortality.

The same result might follow from the combined effect of secular improvement in mortality and class selection increasing in magnitude with increasing age at entry. The latter phenomenon is not improbable, for the class selection exercised by annuitants may well be less powerful at the younger entry ages.

In order to discover which of the above three explanations (if any) is correct, it would be necessary to subdivide the data still further, e.g. the data applicable to entrants in the years 1900–04 might be examined separately for successive periods of exposure, i.e. 1900–04, 1905–09, etc. In theory, this should serve to isolate the effects of the various factors, but in practice the amount of data in the sub-groups would probably be too small to provide reliable figures. This is, in fact, a good example of the difficulties of achieving homogeneity referred to in Chapter XIII.

It is important to note that the authors of the report did not advance the view that calendar year of entry was the only factor, other than age, which influenced the mortality of annuitants in 1900–20. On various occasions, they referred to the many different factors which were operating, but they reached the conclusion that, on balance, the rate of mortality for either sex could be expressed with sufficient accuracy as a function of two variables only, i.e. age attained and year of entry.

The underlying causes of the progression of mortality rates are of considerable importance when we come to justify the use of projected rates. In particular, projection is open to criticism if we take the view that changes in the degree of class selection have been primarily responsible for the improvement in mortality. The projection of mortality rates is justifiable if we consider that the improvement is due to an influence having a definite trend which seems likely, on general grounds, to continue in the future. It is quite another matter if the strongest influence making for mortality improvement is largely fortuitous in its direction and intensity. Over a period of years, conditions might have been such as to bring about a steady improvement in the type of lives effecting annuities, but there is no reason why this trend should persist in

the future. In the circumstances, any attempt to forecast mortality would be open to criticism, for we can hardly hope to be successful if the trend of the most important influence other than age and sex cannot be predicted.

We shall not attempt to delve further into these controversial topics. In summing up, we shall be on safe ground in saying that, whatever views may be held as to the underlying theories, the $a(m)$ and $a(f)$ tables are convenient practical instruments which make some allowance for the secular improvement in mortality.

6. Government Annuitants' Experience 1900–20: method of projection.

Mortality rates were projected horizontally for decennial ages and quinquennial calendar years from 1928 to 1968. Annuity values applicable to entrants in 1928 were calculated therefrom at two rates of interest and these values were compared with the corresponding values based on the mortality rates experienced in the years 1900–20. It was decided that, for entrants in 1928, a suitable allowance would be made for future improvement in mortality by adding 3 per cent for males and 4 per cent for females to the annuity values based on the 1900–20 experience rates. No mortality table based on forecast rates seems to have been constructed.

As in the case of the office experience, it was found that the select rates failed to run into the ultimate after five years. This was, however, attributed to initial temporary selection rather than to spurious selection. For convenience, select rates for the first year only were used, small adjustments being made in the annuity values at the older entry ages to correct the resulting errors.

The main difference between the methods used for the Offices' and the Government investigations lies in the procedure adopted to obtain average values applicable to a given year of entry. In the former case, average mortality rates for entrants in 1925 were estimated and hence annuity values were calculated. In the latter case, the average value of $\dfrac{a'_x}{a_x}$ was estimated for entrants in 1928, where a'_x is the annuity value based on forecast rates and a_x the

value based on the rates experienced in 1900–20. As we have seen, the average value was taken to be 1·03 for males and 1·04 for females and these ratios were used at all ages at entry and for all rates of interest to obtain forecast annuity values from the experience values. In both cases, therefore, the difficulty of using a series of forecast tables was overcome by a process of averaging, although the functions which were averaged were not the same.

7. Continuous Investigation.

This investigation, which began on 1st January 1921, was the forerunner of the corresponding investigation into the mortality of assured lives and the general procedure as regards collection of data and calculation of rates was much the same in each case. Mortality tables have not been constructed, but the results for the period 1921–37 have been published in *J.I.A.* Vol. LXXI and *T.F.A.* Vol. XVII. It is important to note that the tables in these publications compare the experience of 1921–37 with the expected according to the forecast rates for the calendar years of experience 1925 and 1935, not with the rates applicable on the average to annuitants entering in 1925—i.e. those used in the $a(m)$ and $a(f)$ tables.

8. Late intimation of deaths.

Before leaving the subject of annuitants' experiences, one feature peculiar to such investigations deserves attention. This is the error which may be caused by late intimation of deaths. There is seldom any inducement to intimate deaths promptly, such as exists in the case of assurance policies, and it is not unusual for notification to be withheld until the next instalment of annuity falls due. Provided that the errors in their returns are intimated by the contributing offices when full information is available—this should always be within a year of the date as at which the returns were made, since annuities are never paid less frequently than yearly—no harm will be done, but failing this the mortality rates will of course be understated.

CHAPTER XVIII

COMPARISON OF MORTALITY TABLES

1. Principal objects of comparison.

Whatever our object in constructing a mortality table, we shall almost invariably want to compare our results with those obtained from other experiences. Such a comparison may be made for varying reasons which we can broadly classify as financial and statistical. An example of the former occurs when a new standard table is being prepared. The differences between the new rates of mortality and those by an existing table, though interesting, are less important than the variations in the monetary functions. Similarly, when we are investigating the experience of an individual office or of a pension fund and comparing the results with a standard table, we compare the values of q_x chiefly to discover what adjustment in the standard monetary values will render them suitable for use in the individual office or fund.

Cases may occur, however, where a comparison of the actual mortality rates is the essential part of the enquiry, as, for example, when we are investigating the relative mortality of various groups of lives differing from each other in some important respect. The Registrar-General's enquiry into the variation of mortality according to occupation is an example of this type of comparison on a large scale.

In this chapter we shall consider some of the methods by which such comparisons may be made.

2. Comparison of rates of mortality.

The most obvious method is to examine the values of q_x in the experiences age by age and, as in other fields, the obvious method is in some ways the best. This is the only method which takes into account all the points of difference, but the attention to detail which it involves makes it difficult to obtain a general picture of the relative

values. To overcome this difficulty a number of expedients have been evolved:

(*a*) The construction of a graph of the two curves of q_x. This gives a picture of the general trends of the functions, and, in particular, shows clearly between what limits one set of values is greater than the other. It does not, of course, display arithmetically the actual differences which exist and should be regarded as a supplement to rather than a substitute for the table of q_x.

(*b*) The tabulation of the values of q_x for quinquennial ages. These values should indicate clearly, though less vividly than method (*a*), the relative trends as well as the actual and relative differences in q_x. The method has the disadvantage that, in order to avoid distortion by random errors, graduated values must be used. The comparison cannot therefore be made until a late stage: moreover, a new factor is introduced, i.e. the possibility of mistakes in graduation due to paucity of data or to faulty judgment on the part of the actuary. Furthermore, while simple statistical tests are available to measure the significance of the crude rates, these tests are not suitable for application to the graduated rates.

(*c*) The tabulation of quinquennial values of ${}_5q_x$ or ${}_5p_x$. When the data for five ages are combined in this way the random errors are likely to be smaller and it may be feasible to use ungraduated functions. The comparison can therefore be made at an earlier stage than under (*b*) and it is easier to measure the reliability of the evidence. For these reasons, ${}_5q_x$ and ${}_5p_x$ are more satisfactory measures than q_x; of the two, ${}_5q_x$ is the more convenient in practice because of its smaller numerical size. If, for example, the two values of ${}_5q_x$ which we are comparing are ·02493 and ·02618, it is easier to gauge the importance of the difference of ·00125 than to compare the corresponding values of ${}_5p_x$, i.e. ·97507 and ·97382.

3. Comparison of expectations of life.

Another function which is sometimes used is the expectation of life, e_x, and this measure has a considerable popular appeal. It is not unusual to hear of the value of e_0 according to a recent national life table being quoted as a measure of the average lifetime of the

population—regardless of the fact that the phrase "average lifetime" is meaningless except when applied to a particular generation, and that the data on which the table was based were drawn from a series of generations experiencing different mortality.

Even in the hands of the actuary who uses e_0 purely as a composite measure of the rates of mortality experienced during a certain period by a certain body of lives, the function is of limited value, for a comparison of two values of e_0 gives no indication of the relative values of the rates of mortality at individual ages or groups of ages.

A more satisfactory comparison is obtained by considering not only e_0 but a series of values of e_x for, say, decennial ages. This is not a suitable way of displaying the incidence of the difference in mortality, for the fact that $\frac{e_{x+10}}{e'_{x+10}} > \frac{e_x}{e'_x}$ gives only a vague idea of the ratio of q_x to q'_x between ages x and $x+10$. It is, however, useful when we are concerned less with the relation between the rates at individual ages or groups of ages than with the composite effect of the differences at all ages beyond age x. In the case of lives purchasing annuities or effecting whole-life policies, for example, we are concerned with the whole lifetime subsequent to the age at entry, and a comparison of the values of e_x according to two tables is quite a good guide to the relative financial effects of using these tables. On the other hand, if our main object is to compare the financial effects we should be well advised to take a little extra trouble and calculate specimen values of the appropriate monetary functions, e.g. annuity values or rates of premium.

4. Comparison of actual and expected deaths.

An alternative method, which allows a comparison to be made between a new experience and an existing table without calculating values of q_x for the former, is to tabulate in quinquennial age groups the number of deaths which would have occurred if the values of q_x according to the existing table had been experienced and to compare them with the number of deaths that actually occurred. This is known as a comparison of the actual deaths with the expected deaths according to a particular table. The procedure is

shown below, the accented symbols denoting figures based on the old table.

Age	E_x	q'_x	$\theta'_x = E_x \times q'_x$ (expected deaths)	θ_x (actual deaths)
20	403	·00215	·87	2
21	791	·00219	1·73	1
22	1107	·00224	2·48 } 14·0	1 } 11
23	1638	·00230	3·76	4
24	2196	·00236	5·18	3
...	...	...	...	...
...	...	...	...	...
...	...	...	...	...

Central age of group	Expected deaths	Actual deaths	Difference +	Difference −
22	14·0	11		3·0
27	19·7	13		6·7
32	30·1	28		2·1
37	38·4	31		7·4
42	51·7	53	1·3	
47	70·0	64		6·0
52	93·8	99	5·2	
57	129·1	142	12·9	
62	168·6	172	3·4	
	615·4	613	+22·8	−25·2
			= −2·4	

The column of differences presents a clear picture of the relative trends of the rate of mortality. In the example, it is apparent that the new experience shows lighter mortality than the old up to about age 45 and heavier thereafter, subject of course to the statistical tests indicating that the differences are unlikely to be due to random errors.

The example also shows the danger of basing any conclusions on the difference between the total actual and expected deaths over the whole range of ages. This difference is very small and suggests that the mortality according to the two experiences is roughly the same, whereas examination of the deaths by age groups shows that this is not the case.

The over-all difference varies not only with variations in the rates but also with changes in the weights attached to them. In the above example the difference would be greater if the data at the younger ages were increased without any variation in the rates of mortality. The over-all difference is not therefore a satisfactory measure unless the ratio of the rates of mortality according to the two experiences is roughly the same in all age groups. The question of weighting is one of great importance in the study of comparative measures and will be discussed again later in the chapter.

5. Mortality index figures.

We must now consider how a comparison can be carried out when a number of sets of rates of mortality have to be examined. To fix our ideas, let us suppose that the results of a comprehensive investigation into occupational mortality are under review. In these circumstances, the measures of comparison described in paragraph 2 are rather cumbersome, and it is an advantage to have a single composite measure or index of mortality which will enable the statistician to make a rough over-all comparison of the mortality in the various classified groups before he examines the position in more detail. We must not expect such an index to do more than give a general indication of the mortality experienced. Its usefulness will be strictly limited, and it may even be a dangerous weapon if its limitations are disregarded.

With those points in mind, let us consider the various composite measures or indices which have been used:

(*a*) *The crude death-rate.*

This is the ratio of deaths to population for all ages combined and should not be confused with the crude rate of mortality, i.e. the ungraduated rate. It is frequently quoted in making registration returns for towns and districts; its value as a comparative measure is, however, severely restricted since it depends on the age distribution of the population as well as on the mortality rates at individual ages. In the case of an individual town, the distribution of whose population has remained virtually unchanged over a number of years, the crude rate affords a rough measure of the

improvement or deterioration in mortality from year to year, but is useless as a means of comparing the mortality of one town with that of another unless the age distributions are similar. A good example of the fallacious conclusions which may be drawn is that of a health resort which experiences a heavier crude death-rate than the country as a whole, as a result of the exceptionally high proportion of elderly people living there.

As a measure of occupational mortality, this function would be of little value, for the age distribution in different occupations varies considerably, e.g. the number of doctors between the ages of 20 and 25 is a much smaller proportion of the number between ages 20 and 65 than is the case for manual workers such as miners or labourers.

(*b*) *The expectation of life.*

The disadvantage of employing the expectation of life or, in the case of a comparison covering a limited range of ages, the temporary expectation of life, e.g. $e_{20:\overline{45}|}$, is that it is not a simple function of the rates of mortality at individual ages or groups of ages. The significance of any differences between the values of the function for a number of classified groups is therefore more difficult to assess than in the case of some of the measures dealt with later in this paragraph.

The expectation of life is more susceptible to changes in the rate of mortality at young than at old ages, and for this reason some statisticians prefer it to any other index when occupational mortality is under review. They argue that in such an investigation we are mainly concerned with the effect of mortality on the length of the working lifetime of each occupational group, and that more weight should therefore be attached to the rate of mortality at the young than at the old ages.

This is a valid argument, but it loses a good deal of its force when we remember that the mortality index is only intended for use as a rough guide to the general level of mortality in the classified groups preliminary to a more detailed study of age groups. Moreover, an investigation into occupational mortality does not necessarily have for its sole object the measurement of the effect of

mortality on the length of the working lifetime of the different groups, and one is therefore reluctant to adopt a general index which attaches much more weight to the rates of mortality at some ages than at others, regardless of the age distribution of either the occupational group or of the general population.

Much depends on the preference of the individual, but in view of the disadvantage mentioned at the beginning of this paragraph, the use of e_x or $e_{x:\overline{n}|}$ as a mortality index is not generally favoured.

(c) *The ratio of actual to expected deaths.*

In order to obtain a fair comparison we must use the same table to calculate the expected deaths in each classified group. Up to a point, it does not greatly matter what table is chosen, but as a rule it is best to use a suitable standard table, preferably one based on the rates experienced by the combined data of the classified groups. If, therefore, we were investigating the mortality of the male population classified according to occupation, we would use a mortality table based on the experience of the male population as a whole, i.e. a national life table.

Let m_x represent the central rate of mortality at age x according to the standard table and m_x^a and P_x^a the central rate at age x and the population at age x last birthday in a particular classified group. The ratio can then be expressed in symbols as follows:

$$\frac{\Sigma m_x^a P_x^a}{\Sigma m_x P_x^a}. \qquad \text{......(1)}$$

m_x is used instead of q_x because the population aged x last birthday is an approximation to the exposed to risk in central form. If desired, the ratio may be multiplied by a constant such as 100 or 1000 to avoid fractional values. This constant gives the basic value of the index, i.e. the value obtained when $m_x^a = m_x$ at all ages.

Some of the possible defects of this index were mentioned in paragraph 4.

(d) *Comparative mortality figure* (C.M.F.).

Another measure of comparison is obtained if we apply the actual and standard rates of mortality, not to the actual population of the classified group as in the comparison of actual and expected

deaths, but to a standard population. Thus if P_x be the standard population at age x l.b.d. and θ_x the deaths at age x l.b.d. on the basis of the standard table, the C.M.F. for a group whose central rate of mortality at age x is m_x^a is given by

$$\frac{\Sigma m_x^a P_x}{\Sigma m_x P_x} = \frac{\Sigma m_x^a P_x}{\Sigma \theta_x}. \qquad \ldots\ldots(2)$$

For convenience, the ratio is usually expressed as a percentage (or occasionally per mille). Expressed in words, the C.M.F. is the number of deaths which would occur in the standard population if it were subject to the mortality of the classified group, the size of the standard population being such as to produce 100 deaths on the basis of the standard mortality.

(*e*) *Standardized death-rates.*

Instead of using indices (*c*) and (*d*) for comparative purposes, some operators prefer to compare the ratios obtained by multiplying the crude death-rate by one or other of these indices. As the crude death-rate is constant for all classified groups, the comparison differs only in form from that obtained by using the corresponding index. The ratio is known as a *standardized death-rate* and the index is the *standardizing factor.*

When the C.M.F. is used for this purpose, the resulting standardized death-rate is

$$\frac{\Sigma m_x^a P_x}{\Sigma m_x P_x} \times \frac{\Sigma m_x P_x}{\Sigma P_x} = \frac{\Sigma m_x^a P_x}{\Sigma P_x}, \qquad \ldots\ldots(3)$$

i.e. the crude death-rate which would be experienced by the standard population if it were subject to the mortality of the classified group. It should be noted that for this purpose the C.M.F. is based on unity and not on 100 or 1000.

Where the ratio of actual to expected deaths is used as the standardizing factor, the resulting rate is

$$\frac{\Sigma m_x^a P_x^a}{\Sigma m_x P_x^a} \times \frac{\Sigma m_x P_x}{\Sigma P_x}, \qquad \ldots\ldots(4)$$

which cannot be reduced to a simple form. It has, nevertheless, been put to a useful purpose by the Registrar-General for comparing the mortality experience of different towns and districts

between censuses when populations in age groups are unobtainable. For this purpose, the rate is expressed as follows:

$$\frac{\Sigma m_x^a P_x^a}{\Sigma P_x^a} \times \frac{\dfrac{\Sigma m_x P_x}{\Sigma P_x}}{\dfrac{\Sigma m_x P_x^a}{\Sigma P_x^a}}, \qquad \text{......(5)}$$

i.e. the crude death-rate of the classified group multiplied by an adjusting factor. This crude death-rate is obtainable during an intercensal period. The adjusting factor is calculated at the time of the last census and it is assumed that it will remain roughly constant during the intercensal period. This is likely to be the case unless there have been major movements of population such as occur in wartime.

Expression (5) is a good example of an expedient which may have to be devised to obtain a comparison.

The use of a standardized death-rate based on a standard population, as in expression (3), is sometimes called the direct method of standardization. When, however, standard rates of mortality are used in conjunction with the age distribution of the classified group population, as in expression (4), this is known as the indirect method. In the same way, the C.M.F. and the ratio of actual to expected deaths are sometimes called the mortality indices according to the direct and indirect methods respectively.

6. Comparison of mortality indices.

The indices (*c*) and (*d*) of the last paragraph were both used in the Registrar-General's 1921 and 1931 reports on occupational mortality, and are now the generally accepted indices for use when a large number of sets of rates have to be compared. It is therefore worth while examining in more detail the difference between them.

This difference is clearly a matter of weighting. In calculating the ratio of the actual to the expected deaths we weight the actual and the standard rates of mortality with the population in the classified group: in calculating the C.M.F. we weight the rates of

mortality with the numbers in the standard population. The C.M.F. has therefore two advantages:

(i) It provides a constant system of weights for all the classified groups: from the purely theoretical standpoint this is an essential feature of any index by which classified groups are compared with each other, as opposed to a measure used solely for comparing an individual group with the standard table.

(ii) The index can be regarded as a number of deaths instead of a ratio (see paragraph 5 (*d*)) and can therefore be split up into the figures for any desired age groups. The following table illustrates the point:

Age x	Standard population			Group I		Group II	
	P_x	m_x	$m_x P_x$	m_x^a	$m_x^a P_x$	m_x^b	$m_x^b P_x$
30	10,000	·0020	20	·0025	25·0	·0015	15·0
45	15,000	·0040	60	·0035	52·5	·0040	60·0
60	3,000	·0066	20	·0073	21·9	·0087	26·1
Total	28,000		100		99·4		101·1

C.M.F. Group I = 99·4; C.M.F. Group II = 101·1.

The figures for the various classified groups in any particular age group, e.g. 25·0, 15·0, etc., for age 30, form of themselves a consistent series of index numbers, which can be used to compare the mortality in that age group.

On the other hand, a table setting out the expected deaths in age groups for the various classified groups would be of no value since the figures for the different classified groups relate to different populations. To obtain a proper comparison, it would be necessary to calculate the ratio of actual to expected deaths for each age group in each classified group, and these ratios for a particular classified group would not add up to the ratio for the whole group.

The use of the C.M.F. for separate age groups partially overcomes the principal defect of composite measures of mortality, i.e. the fact that they do not show the incidence of differences in mortality. The comparison obtained in this way is somewhat similar to a comparison of the values of ${}_{10}q_{25}$, ${}_{10}q_{35}$, etc., but the

latter method is less satisfactory in that the corresponding single composite measure is the crude death-rate, which, as we have already seen, is of little value, and also because the function ${}_{10}q_x$ is not additive.

As an offset to these advantages of the C.M.F., the fact that it is based on a constant system of weights introduces a serious objection. It means that the same weight is attached to the experience of a particular age group whether the data in the age group are a large or small proportion of the total data in the classified group. Clearly this is unsatisfactory, because the larger the quantity of data in an age group, the more reliable is the corresponding mortality figure likely to be and the more weight should be attached to it in calculating the mortality index for the complete range of ages. In extreme cases the failure of the C.M.F. to take account of this factor may seriously impair its usefulness. The classic example occurred in the Registrar-General's occupational investigation of 1921, in which the C.M.F. for barristers was 1171 (basic value 1000), while the ratio of actual to expected deaths was 107 per cent, a difference of about 10 per cent. This was due to the fact that a single death occurred in age group 20–24 in which the population was only 102, and this very high and entirely unreliable death-rate had a marked effect on the C.M.F., since the proportion of lives aged 20–24 in the standard population was much greater than in the particular occupation.

Even if the data in each age group are sufficient to produce fairly reliable mortality figures, it may be argued that the weight to be attached to an age group should depend on the quantity of data in that group, on the grounds that the larger the proportion of data in an age group the more is that age group representative of the classified group as a whole.

The ratio of actual to expected deaths, though lacking the advantages of the C.M.F., avoids this defect. It is therefore a more reliable index when the age distributions of some of the classified groups differ considerably from that of the standard population, but otherwise the C.M.F. is preferable. The levels of the two indices relative to their basic values will usually be approximately the same, and, as any such index is at best only a rough measure,

the choice of method is in general largely a matter of personal preference. It is, however, important to be on the look-out for any unusual age distributions when using the C.M.F.

The tabulation of both indices provides a useful danger signal, since in those cases where the C.M.F. is unsatisfactory owing to the presence of an abnormal rate of mortality in a particular age group in which the data are scanty, this will be reflected in the difference in the levels of the two indices.

The following rather extreme example illustrates how differences in the two indices can arise. For simplicity only two age groups have been included:

Age group	20–29	30–39
Standard population	10,000	8000
Standard deaths	45	80
m_x (for central age)	·0045	·0100
Group population	50	100
Group deaths	1	1
m_x^a	·0200	·0100

$$\text{C.M.F.} = \frac{\cdot 02 \times 10{,}000 + \cdot 01 \times 8000}{45 + 80} \times 1000 = 2240.$$

Ratio of actual to expected deaths

$$= \frac{2}{\cdot 0045 \times 50 + \cdot 0100 \times 100} \times 100 = 163.$$

It must not be thought that, even if the relative levels of the two indices are identical, they will necessarily provide a satisfactory picture of the mortality of the classified group relative to that of the standard population. The indices would have the same relative values if the age distribution of the group were exactly proportional to that of the standard population, however much the ratios of the corresponding rates of mortality varied between one age group and another. It cannot, in fact, be emphasized too strongly that a really satisfactory comparison is impossible without either calculating values of ${}_5q_x$ or some similar function, or obtaining values of the C.M.F. for several age groups as well as for the whole range of ages included in the investigation.

7. Special indices devised for particular cases.

The choice of index is sometimes determined by the limitations of the data. The ratio of actual to expected deaths can be calculated if we know the rates of mortality according to the standard table, the age distribution of the classified group and the total number of deaths therein regardless of age, even although the rates of mortality experienced by the group are unknown. The C.M.F., however, requires the rates of mortality in the classified group, the age distribution of the standard population and the total deaths in the standard population but not the standard rates of mortality.

If the data are unsuitable for either of these indices, we may have to devise some other index. Suppose, for example, that we are given the number of deaths in age groups for each of the classified groups and the total population of each classified group regardless of age, as well as a suitable standard table. We could not then calculate the C.M.F. or the expected deaths, but we could calculate the "expected population", i.e. the population which would have given rise to the deaths in the classified group if the standard rates of mortality had been experienced.

The index would be the ratio of the expected to the actual populations, instead of actual to expected, in order that an increase in mortality would involve an increase in the index. Expressed in symbols, the index would be

$$\frac{\Sigma\frac{\theta_x^a}{m_x}}{\Sigma P_x^a}=\frac{\Sigma\frac{m_x^a}{m_x}\times P_x^a}{\Sigma P_x^a}, \qquad \ldots\ldots(6)$$

that is to say, the weighted mean of the ratio $\frac{m_x^a}{m_x}$, taking the populations in the classified group as the weights. It will be noticed that the inherent use of these weights does not require a knowledge of their actual values.

CHAPTER XIX

DOUBLE- AND MULTIPLE-DECREMENT TABLES

1. Double-decrement tables.

In calculating life assurance premiums, we normally assume that the lives will remain assured until the policies become claims by death or by the expiry of the term of the assurance. We know that, in practice, a proportion of contracts will cease by lapse or surrender, and, strictly speaking, allowance should be made for this in our premiums. The assurance factor for a whole-life policy should take into account not only the probability of death at each age combined with the present value of the amount payable at death, but also the probability of withdrawal at each age combined with the present value of the amount payable on withdrawal. In the same way, the annuity factor should be the present value of one per annum ceasing on death or previous withdrawal. The errors introduced will, however, be small, provided that the differences between the office surrender values and the reserve values calculated on the same basis as the premiums are small. The surrender values will, in general, be less than the reserve values, and such errors as occur will be on the side of safety. Ordinarily, therefore, in life assurance work, we need not concern ourselves with the question of withdrawals.

Occasions may, however, arise when the effect of withdrawals must be taken into account, and we then require what is known as a double-decrement table in the following form:

Age x	al_x	ad_x	aw_x

The symbol "a" is inserted before l, d and w to avoid confusion with the corresponding single-decrement tables. al_x accordingly

represents the number of survivors to exact age x who have neither died nor withdrawn before attainment of that age. ad_x is the number of deaths between exact ages x and $x+1$, excluding those among lives who had previously withdrawn, and aw_x is the corresponding number of withdrawals. The functions at successive ages are connected by the relation

$$al_{x+1} = al_x - ad_x - aw_x.$$

2. Types of double-decrement tables.

Double-decrement tables are required in various branches of actuarial work involving forces of decrement other than withdrawal. One of the commonest types of table is the combined marriage and mortality table used to calculate contributions and evaluate benefits in any fund where the benefits depend on marital status, e.g. a widows' pension fund providing a pension to the widow on the death of a married member. For example, in investigating the experience of the bachelor members of such a fund, we require to consider two forces of decrement, marriage and mortality. Similarly, when total and permanent disability benefits are payable under an assurance contract, we must find not only the probability of a healthy or active life dying but also the probability of his becoming disabled. Here our two causes of decrement are death and disablement. We may even have more than two forces of decrement simultaneously at work as, for example, in certain staff pension funds where a member may die, withdraw or retire, the benefits payable varying according as he dies in the service, withdraws from it, or retires in circumstances which would entitle him to a pension in terms of its rules. In such a case we require a multiple-decrement table, and the problems of constructing these are discussed later in the chapter.

The theory of constructing a double-decrement table is in all cases the same; we shall accordingly develop it with special reference to the first type of table mentioned above where forces of death and withdrawal are involved. At a later stage we shall discuss some of the practical problems peculiar to other types of double-decrement tables.

3. Dependent rates of decrement.

As in the case of the ordinary life table, the functions al_x, ad_x, etc. are not obtainable direct from the crude data, but the table is constructed by using the rates of decrement which have been calculated from these data. For this purpose, it is convenient to introduce a new function, the *dependent rate* or *dependent probability of decrement*, which may be defined thus:

"Where a body of lives is subject to certain specific causes of decrement but to no other movements, the dependent rate or dependent probability of a particular cause of decrement at age x is the probability that a life aged x will, between ages x and $x+1$, be removed by that particular cause of decrement."

Suppose that there are P_x lives aged x subject to decrement by death and withdrawal but free from any other movements. If θ_x is the number of deaths occurring before age $x+1$, omitting those which took place after the lives had withdrawn, and w_x the corresponding number of withdrawals, the dependent rates, aq_x^d and aq_x^w, are given by the relations

$$aq_x^d = \frac{\theta_x}{P_x}, \quad aq_x^w = \frac{w_x}{P_x}. \qquad \text{......(1)}$$

The rates that we have discussed in earlier chapters are independent rates, the rate of mortality being independent of the rate of withdrawal and vice versa. Using the above data, the independent rates, q_x^d and q_x^w, are given by the relations

$$q_x^d = \frac{\theta_x}{P_x - \frac{1}{2}w_x}, \quad q_x^w = \frac{w_x}{P_x - \frac{1}{2}\theta_x}, \qquad \text{......(2)}$$

assuming both deaths and withdrawals to be evenly distributed over the year of age—assumptions which will be discussed more fully later in this chapter.

Relations (1) and (2) are misleading in appearance. θ_x/P_x includes no term in w_x, while the value of the expression $\dfrac{\theta_x}{P_x - \frac{1}{2}w_x}$ appears to depend on w_x. Hence the former is sometimes mistaken for the independent and the latter for the dependent rate. A clearer conception of the problem may be obtained by a consideration of the effect of an increase in the number of withdrawals. The exposed

to risk of death, $P_x - \frac{1}{2}w_x$, would be reduced and there would be a proportionate reduction in θ_x, provided that the lives withdrawing experienced the same rate of mortality as the remainder. The value of q_x^d would not therefore be affected. In the case of aq_x^d, however, the reduction in θ_x would not be balanced by any change in the denominator and the value of aq_x^d would therefore be reduced.

The relations connecting the dependent and the independent rates follow at once from (1) and (2):

$$\left.\begin{aligned} q_x^d(1-\tfrac{1}{2}aq_x^w) &= aq_x^d, \\ q_x^w(1-\tfrac{1}{2}aq_x^d) &= aq_x^w. \end{aligned}\right\} \qquad \text{......(3)}$$

In the more complicated case where forces of increment and decrement other than death and withdrawal are operating between ages x and $x+1$, let E_x^d and E_x^w represent the exposed to risk of death and withdrawal respectively at age x, and let θ_x and w_x be the number of deaths and withdrawals between ages x and $x+1$. Then

$$q_x^d = \frac{\theta_x}{E_x^d}, \quad q_x^w = \frac{w_x}{E_x^w}. \qquad \text{......(4)}$$

Let aq_x be the probability of decrement by either death or withdrawal, i.e. the probability that a life aged x will either die or withdraw before age $x+1$ when no other forces are operating. The numerator of aq_x is $\theta_x + w_x$ and the denominator will be the exposed to risk obtained by treating both deaths and withdrawals as exposed for a full year at age x. Denoting this denominator by E_x, we have

$$aq_x = \frac{\theta_x + w_x}{E_x}.$$

This total probability can be split up into its component parts, i.e. the dependent probabilities of death and withdrawal, and we have

$$aq_x^d = \frac{\theta_x}{E_x}, \quad aq_x^w = \frac{w_x}{E_x}, \quad aq_x = aq_x^d + aq_x^w. \qquad \text{......(5)}$$

Now from our definition of E_x, it follows that

$$E_x = E_x^d + \tfrac{1}{2}w_x = E_x^w + \tfrac{1}{2}\theta_x. \qquad \text{......(6)}$$

Hence

$$\left.\begin{aligned} q_x^d &= \frac{\theta_x}{E_x^d} = \frac{\theta_x}{E_x - \frac{1}{2}w_x}, \\ q_x^w &= \frac{w_x}{E_x^w} = \frac{w_x}{E_x - \frac{1}{2}\theta_x}. \end{aligned}\right\} \qquad \text{......(7)}$$

Relations (5) and (7), which are analogous to (1) and (2), lead to the same relations as before between the dependent and the independent rates. It should be noted that the arguments of this paragraph apply whether x denotes the exact age or merely the assumed age.

If the rates of death and withdrawal according to a particular experience are to be used to construct a double-decrement table without any modification except by graduation, there is no need to calculate the independent rates. The dependent rates can be obtained direct from the crude data, graduated and applied to construct the required table by fixing a radix and proceeding on the usual lines.

Sometimes, however, the independent rates may be required. If, for example, we wished to compare the results with those of other investigations, the independent rates should be used, for in comparing rates of mortality we should as far as possible eliminate the effect of the rates of withdrawal; these might vary considerably between two investigations and, as we have already seen, a variation in the rate of withdrawal affects the dependent but not the independent rate of mortality.

Again, it may be desired to base a double-decrement table on the rates of withdrawal according to the experience under review and the rates of mortality according to a standard mortality table. For this purpose the independent rates of withdrawal should be calculated. After graduation, these rates, together with the standard rates of mortality, would be substituted in formulae (3) and the equations solved to obtain the corresponding dependent rates from which the double-decrement table could be constructed.

4. Relations between dependent and independent rates of decrement.

To examine further the relations between the various rates of decrement, let us represent the functions with which we are concerned by the symbols in the following table, the particular description having no meaning in those cases where blanks appear.

Description of function	Type of decrement		
	Death	Withdrawal	Both together
Force of decrement	μ_x^d	μ_x^w	$a\mu_x$
Dependent rate of decrement	aq_x^d	aq_x^w	aq_x
Independent rate of decrement	q_x^d	q_x^w	
Dependent probability of a life aged x surviving to age $x+t$			${}_tap_x$
Independent probability of a life aged x surviving to age $x+t$	${}_tp_x^d$	${}_tp_x^w$	

The forces of death and withdrawal at age x operate at that particular point of age only and clearly are not interdependent. There cannot, therefore, be a dependent force of decrement in the same sense as there is a dependent rate of decrement. The three forces in the table are connected by the relation

$$a\mu_x = \mu_x^d + \mu_x^w. \qquad \ldots\ldots(8)$$

The rates of decrement may be expressed as follows:

$$\left.\begin{aligned}
aq_x^d &= \int_0^1 {}_tap_x\,\mu_{x+t}^d\,dt, \quad aq_x^w = \int_0^1 {}_tap_x\,\mu_{x+t}^w\,dt,\\
q_x^d &= \int_0^1 {}_tp_x^d\,\mu_{x+t}^d\,dt, \quad q_x^w = \int_0^1 {}_tp_x^w\,\mu_{x+t}^w\,dt,\\
aq_x &= \int_0^1 {}_tap_x\,a\mu_{x+t}\,dt = \int_0^1 {}_tap_x(\mu_{x+t}^d + \mu_{x+t}^w)\,dt = aq_x^d + aq_x^w.
\end{aligned}\right\} \qquad \ldots\ldots(9)$$

The last relation in (9) could of course be obtained more directly by reference to the double-decrement table. Again,

$$\begin{aligned}
{}_tap_x &= \exp\left[-\int_0^t a\mu_{x+s}\,ds\right]\\
&= \exp\left[-\int_0^t \mu_{x+s}^d\,ds - \int_0^t \mu_{x+s}^w\,ds\right]\\
&= {}_tp_x^d \times {}_tp_x^w.
\end{aligned} \qquad \ldots\ldots(10)$$

Hence $$aq_x^d + aq_x^w = aq_x = 1 - ap_x$$
$$= 1 - p_x^d \times p_x^w$$
$$= 1 - (1 - q_x^d)(1 - q_x^w). \qquad \text{......(11)}$$

5. Distribution of deaths and withdrawals.

Let us now examine more closely the assumptions made in relations (2). In the case of q_x^d, by excluding $\frac{1}{2}w_x$ in the denominator we assumed in effect that the actual withdrawals were uniformly distributed over the year of age, i.e. when the data were subject to the particular rate of mortality which was experienced. Obviously, if the assumption were justified in these conditions, it would be incorrect if the same rate of withdrawal operated in conjunction with a different rate of mortality and, in particular, it would not apply in the special case where withdrawal was the only cause of decrement. (This is, of course, a purely theoretical concept, for it is impossible in practice to have a body of lives not liable to decrement by death—a fact which does not, however, detract from the value of the independent rate of withdrawal as a comparative measure.) As regards the deaths, we have, in effect, assumed that they would be evenly distributed if no other causes of decrement were operating, for this assumption is necessary to justify the inclusion in the ratio q_x^d of fractional exposures and the corresponding deaths (see II, 5).

In the case of q_x^w, on the other hand, the implied assumptions are that the withdrawals would be evenly distributed if there were no other causes of decrement, and the deaths would be evenly distributed if the data were subject to the rate of withdrawal actually experienced. The implied assumptions for q_x^d and q_x^w are therefore inconsistent, and we must enquire whether any serious practical error has been introduced thereby. The simplest way to do this is to start with consistent and sufficiently correct assumptions regarding the distribution of deaths and withdrawals and see what difference this will make to our results. Let us assume that in each case the distribution is uniform when the force in question is the only force of decrement operating between ages x and $x+1$. In the notation of para. 4 we then have

$${}_tp_x^d \mu_{x+t}^d = k, \quad {}_tp_x^w \mu_{x+t}^w = l$$

where k and l are constants $0 \leqslant t \leqslant 1$. Then

$$q_x^d = \int_0^1 k\,dt = k,$$

$$q_x^w = \int_0^1 l\,dt = l,$$

$$aq_x^d = \int_0^1 {}_t ap_x \mu_{x+t}^d\,dt = \int_0^1 {}_tp_x^d \times {}_tp_x^w \mu_{x+t}^d\,dt$$

$$= \int_0^1 k\,{}_tp_x^w\,dt.$$

But
$${}_tp_x^w = 1 - \int_0^t {}_sp_x^w \mu_{x+s}^w\,ds = 1 - \int_0^t l\,ds$$
$$= 1 - lt.$$

Hence
$$aq_x^d = \int_0^1 k(1-lt)\,dt = k(1-\tfrac{1}{2}l)$$

and
$$\left.\begin{aligned} &= q_x^d(1-\tfrac{1}{2}q_x^w),\\ aq_x^w &= q_x^w(1-\tfrac{1}{2}q_x^d).\end{aligned}\right\} \qquad \ldots\ldots(12)$$

Although these relations are more consistent than relations (3) they are on the whole less convenient, for the calculation of the independent from the dependent rates requires the solution of quadratic equations.

Relations (3) give the following values of the dependent in terms of the independent rates:

$$\left.\begin{aligned} aq_x^d &= \frac{q_x^d(1-\frac{1}{2}q_x^w)}{1-\frac{1}{4}q_x^d q_x^w},\\ aq_x^w &= \frac{q_x^w(1-\frac{1}{2}q_x^d)}{1-\frac{1}{4}q_x^d q_x^w}.\end{aligned}\right\} \qquad \ldots\ldots(13)$$

These values are greater than those given by (12) in the ratio of $1 : 1-\frac{1}{4}q_x^d q_x^w$, and it is clear that both rates of decrement must be large to make this ratio significantly different from unity. Even if q_x^d and q_x^w were as large as $\cdot 2$ and $\cdot 5$ respectively, the ratio would be $1 : \cdot 975$, i.e. a $2\frac{1}{2}$ per cent error; under these conditions the assumption as to the distribution of exits over the year of age would probably be far from the truth and the whole basis of the argument would be vitiated.

It appears therefore that our original assumptions, though not strictly consistent in themselves, have led us to relations not substantially different from and for practical purposes more convenient than those reached when assumptions are made which are more satisfactory from the theoretical standpoint.

The assumption that the deaths are uniformly distributed over the year of age is, as mentioned in earlier chapters, an arbitrary one, although in general it is sufficiently close to the truth for practical purposes. The same is not necessarily true of the withdrawals, and circumstances may be such that the withdrawals at age x are exposed on the average for a fractional period differing appreciably from $\frac{1}{2}$. This is particularly liable to occur when the method of grouping the data is according to the assumed age on the policy anniversary instead of by exact ages. Allowance could, however, be made for this by substituting the appropriate fraction for $\frac{1}{2}$ in the denominator of (3). If, for example, w_x were on the average exposed for a fraction h of a year between assumed ages x and $x+1$, we should have

$$q_x^d = \frac{aq_x^d}{1-(1-h)\,aq_x^w}. \qquad \text{......(14)}$$

6. Central rates of decrement.

By the method which led to (3), the central rates of decrement m_x^d and m_x^w may be expressed in terms of the dependent rates as follows:

$$m_x^d = \frac{aq_x^d}{1-\frac{1}{2}aq_x^d-\frac{1}{2}aq_x^w}, \text{ etc.,} \qquad \text{......(15)}$$

whence

$$aq_x^d = \frac{m_x^d}{1+\frac{1}{2}m_x^d+\frac{1}{2}m_x^w}, \text{ etc.} \qquad \text{......(16)}$$

The corresponding expression based on the assumptions underlying relation (12) is

$$aq_x^d = \frac{m_x^d}{(1+\frac{1}{2}m_x^d)\,(1+\frac{1}{2}m_x^w)}, \qquad \text{......(17)}$$

and once again the difference is for practical purposes insignificant. This expression is obtained by using the relations

$$q_x^d = \frac{m_x^d}{1+\frac{1}{2}m_x^d}, \text{ etc.}$$

7. Marriage and mortality table.

This table takes the following form:

Age x	bl_x	bd_x	bm_x	ml_x	md_x

The first four columns form a double-decrement table, bl_x being the number of survivors to age x, bd_x the number of bachelors dying and bm_x the number marrying between ages x and $x+1$. The method of construction follows the usual lines. These columns permit of the building up of commutation functions, from which the value of a unit payable on the death or marriage of a bachelor may be evaluated.

The last two columns represent an extension of the original conception of double-decrement tables. They show how many of the bachelors at the youngest age would survive to age x as married men and how many would die between ages x and $x+1$. The ml_x column is subject to increase by the marriage of bachelors and to reduction by the death of married men. The numbers at successive ages are therefore connected by the relation

$$ml_{x+1} = ml_x + bm_x - md_x. \qquad \ldots\ldots(18)$$

The last three columns are therefore akin to a double-decrement table, but one of the forces is a force of increment. The only additional function to be calculated from the crude data to enable us to construct the last two columns is the rate of mortality among married men. We might of course assume that the rates would be the same as for bachelors, but this is unsound, since marriage is a selective influence and there is more chance of a healthy man marrying than of an under-average life. Alternatively, we might make our investigation into the mortality of the members of the fund regardless of marital status as we should do if we were investigating the mortality experience of a life office. The two cases are hardly on a par, however, for in the former case the benefits depend on marital status, and any differences in mortality are more

important. Whenever possible, therefore, it is desirable that the rates for bachelors and married men should be calculated separately from the data of the fund that we are investigating.

Having obtained the rates of mortality among married men, which we shall denote by q_x^{md}, we can complete the last two columns of the table by successively applying relation (18) and the following relation,

$$md_x = q_x^{md}(ml_x + \tfrac{1}{2}bm_x), \qquad \text{......(19)}$$

the assumption being made that the marriages at age x are evenly spread over the year of age. At the youngest age at which marriages occur, say y, these relations will take the form

$$md_y = q_y^{md} \times \tfrac{1}{2}bm_y, \quad ml_{y+1} = bm_y - md_y,$$

for ml_y is zero.

The above form of table is not really of much practical value for calculating functions relating to married men. If, for example, we required the respective probabilities of a bachelor and a married man aged x dying between ages $x+t$ and $x+t+1$, the bachelor having married in the meantime, these would not be given by the ratios $\dfrac{md_{x+t}}{bl_x}$ and $\dfrac{md_{x+t}}{ml_x}$ respectively, for some of the deaths included in the numerator would relate to lives who married before age x and the remainder to lives who married thereafter; we require only the latter deaths for the probability applicable to a bachelor aged x and only the former deaths for that applicable to a married man aged x. The ratio $\dfrac{md_{x+t}}{bl_x}$ is correct only in the case where $x \leqslant y$, and the ratio $\dfrac{md_{x+t}}{ml_x}$ is correct only if no marriages occur after age x.

The difficulty can be overcome by constructing in addition an ordinary life table for married men, based on the same rates of mortality as were the last two columns of the marriage and mortality table. Let us represent the number of survivors to age x by l_x^{md} and the number of deaths between ages x and $x+1$ by d_x^m, where

$$q_x^{md} = \frac{d_x^m}{l_x^{md}} \quad \text{and} \quad l_{x+1}^{md} = l_x^{md} - d_x^m.$$

This supplementary table gives directly the probabilities of a

married man aged x surviving to age $x+t$ and of his dying between ages $x+t$ and $x+t+1$ if we assume that the rate of mortality among married men does not vary appreciably according to duration since marriage.

To obtain the probability of a bachelor aged x surviving to age $x+t$ as a married man, we must exclude from ml_{x+t} the lives who were married before age x, i.e. $ml_x \times \frac{l^{md}_{x+t}}{l^{md}_x}$, so that the required probability is

$$\frac{ml_{x+t} - ml_x \times \frac{l^{md}_{x+t}}{l^{md}_x}}{bl_x}. \qquad \text{......(20)}$$

A similar expression can be devised for the probability of a bachelor aged x dying a married man between ages $x+t$ and $x+t+1$, namely,

$$\frac{md_{x+t} - ml_x \times \frac{d^{m}_{x+t}}{l^{md}_x}}{bl_x}. \qquad \text{......(21)}$$

Monetary functions can be calculated on similar lines. For example, the value of a unit payable on the death of a bachelor now aged x, if he should be married when he died, would be

$$\frac{1}{bl_x} \sum_{t=0} v^{t+\frac{1}{2}} \left(md_{x+t} - ml_x \times \frac{d^{m}_{x+t}}{l^{md}_x} \right). \qquad \text{......(22)}$$

An alternative method is to construct a table of $\bar{A}^{md}_x$, i.e. the assurance factor based on the table for married men. The required value would then be

$$\frac{1}{bl_x} \sum_{t=0} v^{t+\frac{1}{2}}\, bm_{x+t}\, \bar{A}^{md}_{x+t+\frac{1}{2}}. \qquad \text{......(23)}$$

No mention has been made so far of the contingency of widowerhood and we have treated all married men as if they remained in that category until death. In the case of a widows' fund this would not be satisfactory, as the benefits payable on death would not be the same for a widower as for a married man. One method of allowing for this factor would be to make the married men's life table into a double-decrement table, the causes of decrement being the death of the man and the death of his wife. Problems of this

kind are, however, outside the scope of this book and will be encountered in the study of pension funds.

8. Total and permanent disability table.

The decrements in this case are death and disability. If the benefits are payable in a lump sum or in instalments over a short period, the required monetary functions can easily be calculated from this table. If, however, they are payable in the form of an annuity for the rest of life, account must be taken of the mortality of the disabled lives.

A table analogous to a marriage and mortality table could be constructed on the following lines. Active lives are those who are not disabled, and ai_x represents the number of active lives becoming disabled between ages x and $x+1$:

Age x	Active lives			Disabled lives	
	al_x	ad_x	ai_x	il_x	id_x

This table has the same disadvantage as a marriage and mortality table, however, and, in addition, it makes no allowance for the existence of a considerable amount of temporary initial selection in a reverse direction in the two or three years following disablement. It is therefore desirable to value the benefits payable on disablement by an expression similar to (23), in which the place of the assurance factor would be taken by $a^{id}_{x+t+\frac{1}{2}}$, i.e. the value of an annuity to a life aged $x+t+\frac{1}{2}$ who has just become disabled. $a^{id}_{x+t+\frac{1}{2}}$ would be based on a select mortality table of disabled lives, in which the age at disablement would take the place of the age at entry in an ordinary select mortality table.

9. Multiple-decrement tables.

All that has been written in the preceding paragraphs regarding double-decrement tables can be readily extended to the case where three or more causes of decrement are at work. If we denote the

three forces of decrement by the symbols d, w and r, relations (3) of para. 3 become

$$\left.\begin{aligned} q_x^d\left(1-\tfrac{1}{2}aq_x^w-\tfrac{1}{2}aq_x^r\right)&=aq_x^d,\\ q_x^w\left(1-\tfrac{1}{2}aq_x^r-\tfrac{1}{2}aq_x^d\right)&=aq_x^w,\\ q_x^r\left(1-\tfrac{1}{2}aq_x^d-\tfrac{1}{2}aq_x^w\right)&=aq_x^r,\end{aligned}\right\} \qquad \ldots\ldots(24)$$

and the other relations between the various functions can be developed in the same way as before.

10. Service tables.

The commonest type of multiple-decrement table is that known as a service table. This is used to value the benefits payable under a staff pension fund. We shall discuss it in some detail.

The decrements among the active lives are deaths and retirements and, in many cases, withdrawals. The mortality of the pensioners must also be taken into account, and the problem and its solution are therefore much the same as in the case of a combined marriage and mortality experience. There is this difference, however, that there are not normally any retirements until the older ages, and the value of the benefits can therefore be obtained by means of a table on the lines of that in para. 7 without any supplementary table except for a small proportion of the lives, i.e. those older than the lowest age at which retirement on pension is permissible. For these lives, the benefits must be valued by formulae on the lines of either (22) or (23), suitably adjusted to allow for the fact that the benefit is an annuity instead of an assurance.

Members of a pension fund are, in many cases, permitted to retire before the normal age of superannuation on grounds of ill-health. In that event, normal retirements and those due to ill-health should be recorded separately and treated as separate decrements in the table if the numbers are sufficient to permit of this procedure. Selection at the date of retirement will have to be taken into account for the latter class, and the problem and its solution will be the same as in the case of a permanent disability experience. In these circumstances it may be simpler to deal with normal and ill-health retirements on the same lines and value both benefits by a formula

similar to (23). If ar_{x+t} and ar'_{x+t} represent the number of retirements according to the table between ages $x+t$ and $x+t+1$ and $a^{rd}_{x+t+\frac{1}{2}}$ and $a^{r'd}_{x+t+\frac{1}{2}}$ the values of pensions of one per annum to lives retiring between ages $x+t$ and $x+t+1$, where the dashes signify retirements due to ill-health, the value of a pension of one per annum to an active life now aged x would be

$$\sum_{t=0} v^{t+\frac{1}{2}}\left(\frac{ar_{x+t}}{al_x}\times a^{rd}_{x+t+\frac{1}{2}}+\frac{ar'_{x+t}}{al_x}\times a^{r'd}_{x+t+\frac{1}{2}}\right). \quad \ldots\ldots(25)$$

Before constructing a service table, the rules of the pension fund should be carefully studied to see whether there are any special features which would have to be taken into account. In particular, the rules fixing the ages at which members may or must retire should be examined. If, for example, the minimum and maximum ages of retirement on grounds of age were 60 and 65 respectively and members were in general allowed to retire at their option at any age between these limits, there would be a large number of retirements at or near exact ages 60 and 65 and the remaining retirements would probably be distributed without any concentrations at particular points of age.

In formula (25) we assumed that the retirements between ages x and $x+1$ would be uniformly distributed over the year of age, but in the above example this would not be correct for ar_{60} and ar_{64}. It would therefore be desirable to set up a separate column in the service table for retirements in the neighbourhood of exact ages 60 and 65. For this purpose, when grouping the crude data, we should have to segregate the retirements within a few weeks after these birthdays.

Let $\rho^{(1)}_{60}$ be the number of retirements at exact age 60 and $\rho^{(2)}_{60}$ the number of the remaining retirements between exact ages 60 and 61, and suppose that the data are suitable for an exact age method. There will be two rates of retirement at age 60 which we shall denote by $aq^{r(1)}_{60}$ and $aq^{r(2)}_{60}$, the numerators being $\rho^{(1)}_{60}$ and $\rho^{(2)}_{60}$ respectively. The exposed to risk for the former rate would be the number of lives who passed through exact age 60 during the investigation, including $\rho^{(1)}_{60}$, and the rate would therefore be of a rather unusual type. The exposed to risk for $aq^{r(2)}_{60}$ would be the

same as for the dependent probability of death and would not include any exposure for the retirements at exact age 60.

Special care is necessary in dealing with beginners and enders. If they are grouped by age last birthday, all the enders at age 60 last birthday but none of the corresponding beginners should contribute a unit to $E_{60}^{r(1)}$. If, however, the grouping is by nearest age, the lives should if possible be regrouped by age last birthday, failing which it would have to be assumed that $\frac{1}{2}b_{60}$ and $\frac{1}{2}e_{60}$ passed through exact age 60 during the investigation.

Having obtained the dependent rates, the following relations would be used to build up the service table at age 60:

$$\begin{aligned}
ar_{60}^{(1)} &= al_{60} \times aq_{60}^{r(1)}, \\
ar_{60}^{(2)} &= (al_{60} - ar_{60}^{(1)})\, aq_{60}^{r(2)}, \\
ad_{60} &= (al_{60} - ar_{60}^{(1)})\, aq_{60}^{d}, \\
al_{61} &= al_{60} - ar_{60}^{(1)} - ar_{60}^{(2)} - ad_{60}.
\end{aligned}$$

11. Graded service tables.

The staff of a large concern is sometimes divided into several grades having different salary scales, the members of the higher grades being recruited from the lower grades. In these conditions, the service table can be graded and in the case of a staff divided into three grades A, B and C, the numbers in grade B will be fed by promotions from grade A and will be subject to decrease by deaths, retirements and promotions to grade C, which in turn will be fed by these promotions and subject to decrease by deaths and retirements. The form of the table is in fact merely an extension of the ordinary service table. Pensioners might have to be dealt with separately for each grade, as the pension payable would no doubt vary according to the grade of the pensioner at the date of retirement. This would necessitate columns for the survivors and deaths among the pensioners of each grade which would be built up in the same way as in an ordinary service table.

A graded service table has the same drawbacks as a marriage and mortality table. The only probabilities which can be obtained from it direct are (a) those which relate entirely to the lowest grade, e.g. the probability of a member of grade A aged x dying at age $x+t$

while still in that grade, and (*b*) certain probabilities applicable to lives in grade A whose present age is below the minimum age at which transfers to grade B occur, e.g. if transfers do not begin before age $x+n$ the table would give the probability of a member in grade A aged x dying at any subsequent age in grade B. If there are only two grades, resort may be had to the methods leading to relations (20)–(23), but if there are more than two grades, the resulting formulae will be complicated. In these circumstances the method which led to formula (23) will usually be more convenient than that which underlies formula (22), but there are other approaches to the problem which are outside the scope of our subject. In this book we are concerned only with the method of constructing a graded service table and with its obvious limitations.

Example 1.

An investigation has been made into the mortality of two groups of lives, similar in all respects except that one group is subject to a special extra risk. If the rates of mortality for each group are known, how would you find the rate of mortality due to the extra risk?

The question implies that an independent measure of the extra risk is required, i.e. one which does not depend on any particular rate of normal mortality, in conjunction with which it is operating. An independent rate such as this, like an independent rate of withdrawal, is a theoretical concept in that it cannot operate in practice on its own. Nevertheless, it is quite appropriate to calculate such a rate, in order that it may be compared with other similar rates calculated from different data which may involve different rates of normal mortality.

Let aq_x be the total rate of mortality for the special group, and let q_x^α be the rate for the normal group and q_x^β the rate of extra mortality which is the unknown quantity.

Then the special group is subject to two forces of decrement, α and β, and from relation (11)

$$aq_x = 1-(1-q_x^\alpha)(1-q_x^\beta).$$

Hence

$$q_x^\beta = \frac{aq_x - q_x^\alpha}{1-q_x^\alpha}.$$

Example 2.

All male members of the staff of a large firm join a staff scheme at exact age 21 and are entitled to death benefits which vary according to marital status. A table is required showing the probability of an entrant at age 21 dying (*a*) as a bachelor and (*b*) as a married man between ages

x and $x+1$. Assuming that the data are set out on cards showing the calendar years of entry, withdrawal, marriage and death, how would you construct the table, neglecting the question of graduation?

We must assume that no information about the exact dates of the movements is obtainable. All the available information relates to calendar years and we shall accordingly group our data according to the assumed age at the beginning of the calendar year of exposure.

21 – (calendar year of entry) + (calendar year of movement) gives the age on the birthday in the calendar year of the movement, i.e. the age next birthday on the 1st January in that year. We shall call this "assumed age x". We shall accordingly take as our exposed to risk at assumed age x the number of years' exposure contributed by lives in the calendar years in which they attained exact age x. Let us assume that the period of the investigation is 1935–39 inclusive.

We shall define the various symbols as follows, distinction being made between bachelors and married men at the date of the movement by accenting the symbols applicable to the latter:

b_x = numbers of beginners of assumed age x on 1st January 1935.
e_x = number of enders of assumed age x on 31st December 1939 (in this case x is $21-(\text{year of entry})+1940$.
n_{21} = number of entrants in the years 1935–39.
θ_x = number of deaths at assumed age x in 1935–39.
w_x = number of withdrawals at assumed age x in 1935–39.
m_x = number of marriages at assumed age x in 1935–39.

For bachelors it will be convenient to calculate the dependent rates of marriage and mortality to facilitate the construction of the table. The b_x beginners are exposed for a full year at assumed age x and the e_x enders are not exposed at all at that age. We shall assume that withdrawal occurs on the average in the middle of the calendar year so that the withdrawals are exposed for half a year at assumed age x and not exposed at all thereafter. As we require the dependent rates θ_x and m_x must both be exposed for a full year at assumed age x. We have, therefore,

$$E_x = E_{x-1} + b_x - e_x - \tfrac{1}{2}(w_{x-1} + w_x) - \theta_{x-1} - m_{x-1},$$

$$bq_x^d = \frac{\theta_x}{E_x}, \quad bq_x^m = \frac{m_x}{E_x}.$$

Assuming that birthdays are uniformly spread over the calendar year, these rates will apply to exact age $x-\tfrac{1}{2}$.

At age 21, the exposed to risk formula takes the form

$$E_{21} = \tfrac{1}{2}n_{21} - \tfrac{1}{2}w_{21},$$

there being no beginners or enders aged 21 next birthday and the assumption being made that the new entrants are evenly spread over the calendar year. We are also assuming that none of the entrants are married men. The resulting rates bq^d_{21} and bq^m_{21} do not refer to age 21 next birthday in the same sense as at other ages, for instead of the exposure relating to the year of age $20\frac{1}{2}$–$21\frac{1}{2}$ on the average, it relates to the half year of age 21–$21\frac{1}{2}$ on the average and might therefore be taken as the approximate rate for exact age $20\frac{3}{4}$, which should relate to the year of age $20\frac{3}{4}$–$21\frac{3}{4}$. Furthermore, the average exposure for the withdrawals w_{21} is taken as zero by the above formula, whereas it would be more correct to allow a small fractional exposure. The error involved is probably negligible as w_{21} will be small relative to n_{21}. At age 22 the exposed to risk formula is

$$E_{22}=E_{21}+b_{22}-e_{22}+\tfrac{1}{2}n_{21}-\tfrac{1}{2}(w_{21}+w_{22})-\theta_{21}-m_{21},$$

and this gives the rates for exact age $21\frac{1}{2}$.

The rates for integral ages will be obtained by interpolation and regard must be had to the special significance of bq^d_{21} and bq^m_{21}. The double-decrement table can then be set up in the usual way.

If there are no marriages at assumed age 21, let y be the lowest assumed age at which they occur. bq^m_y refers to exact age $y-\frac{1}{2}$, and in performing the interpolation we should have to decide whether to make $y-1$ or y the youngest integral age at which the rate of marriage has a value. The choice could be determined by the trend of the values of bq^m_x but would probably be largely arbitrary.

As regards the mortality of the married men the marriages provide a flow of new entrants, and we must expose m_x for half a year at assumed age x and for a full year thereafter. The exposed to risk formula will be

$$E'_x=E'_{x-1}+b'_x-e'_x+\tfrac{1}{2}(m_{x-1}+m_x)-\tfrac{1}{2}(w'_{x-1}+w'_x)-\theta'_{x-1},$$

whence

$$q^{md}_x=\frac{\theta'_x}{E'_x}.$$

At age y, the youngest assumed age at which there are any marriages, the formula is

$$E'_y=\tfrac{1}{2}m_y-\tfrac{1}{2}w'_y,$$

and, while similar considerations apply as for the bachelors at age 21, the position is more complicated, since the marriages do not all occur at exact age y, whereas all the new entrants entered at exact age 21. Moreover, it is not correct to assume an even distribution of m_y over the calendar year, since the force of marriage will be increasing rapidly at age y. In practice the data for married men at assumed age y would be too scanty to give a reliable rate of mortality and arbitrary values would

have to be assigned at this age and probably at several succeeding ages also.

Interpolation must next be performed to obtain rates for integral ages and the lowest age for which a rate is required will of course be that for which the first entry appears in the bm_x column. The final column of the table will then be completed by means of relations (18) and (19).

Example 3.

The following table shows part of the experience of a pension fund of a large combine. The retirement rules provide for full pension on attainment of age 60 or on completion of 40 years' service whichever shall happen later, with the proviso that no member shall continue in service after age 65.

What assumptions would you make regarding these retirements and what further information would you call for before deciding to employ this retirement experience for the construction of a service table for use in testing contributions or in valuation?

Find the ratio of actual deaths in service to expected by A 1924–29 ultimate mortality for the group of ages shown in the table. Using this percentage of the A 1924–29 rates for mortality during service and $a^{(m)}$ rates for mortality after retirement, calculate the probability that a member aged 59 exactly will be alive but retired at age 62.

Table

Age x	Beginners on 1st Jan. 1933 b_x	Enders on 31st Dec. 1937 e_x	Retirements during period ρ_x	Deaths during period θ_x
58	520	590	4	46
59	500	620	4	43
60	120	140	1708	11
61	100	110	120	10
62	86	100	86	12
63	54	66	64	6
64	28	32	42	6
Total	1408	1658	2028	134

You are also given

(*a*) that the information above is tabulated according to age last birthday on the date given or on the happening of the event.

(*b*) that among the retirements scheduled at 60 last birthday 1686

took place within a few weeks of the 60th birthday and that the number of retirements at exact age 65 during the period was 123.

(c) that the dependent probability of retirement between age 58 and 59 was ·0016.

The retirements before age 60 will certainly be due to ill-health; some of those between ages 60 and 65 may also be due to this cause and the number of them should be ascertained. While it is better where possible to deal separately with normal and ill-health retirements, in this case the data are inadequate for the adoption of this method. We should, however, consider, in the light of any evidence which is obtainable from other experiences, whether the proportion of ill-health retirements is abnormal.

As might be expected, there is a large number of retirements near exact age 60 and a fairly steady flow between 60 and 65, the balance taking place at exact age 65. The distribution of these retirements depends mainly on the age at entry, for there will be few retirements, except on account of ill-health, before completion of 40 years' service, when full pension is obtainable. In the past, many entrants were evidently recruited after age 20, and the trend of the entry ages in the past 40 years must be examined to see whether a similar age distribution of retirements is to be expected in future.

The latter part of the question suggests that the result of these enquiries has been satisfactory and that the table can be based on the data of the experience except in the case of pensioners, for whom the data were presumably insufficient to provide reliable rates.

The deaths and retirements are grouped by age last birthday and the exact age method is therefore the most suitable. We shall first obtain E_x, the exposed to risk in the form suitable for calculating dependent rates. The corresponding functions for the independent rates, E_x^d and E_x^r, are easily obtainable therefrom.

The b_x beginners who are aged $x+\frac{1}{2}$ on the average must be exposed for half a year at exact age x and for a full year thereafter; and similarly for e_x. Hence, we have

$$E_x = E_{x-1} + \tfrac{1}{2}(b_{x-1} + b_x) - \tfrac{1}{2}(e_{x-1} + e_x) - \rho_{x-1} - \theta_{x-1}.$$

At age 58 we know that the dependent probability of retirement is ·0016 and that $\rho_{58}=4$; hence $E_{58}=2500$. At age 60 we must divide the retirements into two groups, $\rho_{60}^{(1)}$ who retire at exact age 60 and $\rho_{60}^{(2)}$ who retire between 60 and 61, and in calculating E_{60} we must exclude the former but retain the latter. The exposed to risk formula therefore takes a special form at ages 60 and 61:

$$E_{60} = E_{59} + \tfrac{1}{2}(b_{59} + b_{60}) - \tfrac{1}{2}(e_{59} + e_{60}) - \rho_{59} - \rho_{60}^{(1)} - \theta_{59},$$

$$E_{61} = E_{60} + \tfrac{1}{2}(b_{60} + b_{61}) - \tfrac{1}{2}(e_{60} + e_{61}) - \rho_{60}^{(2)} - \theta_{60}.$$

The following is the working table for calculating E_x.

Age x	$\frac{1}{2}(b_{x-1}+b_x)$	$\frac{1}{2}(e_{x-1}+e_x)$	ρ_{x-1}	θ_{x-1}	E_x
58	—	—	—	—	2500
59	510	605	4	46	2355
60	310	380	1690	43	552
61	110	125	22	11	504
62	93	105	120	10	362
63	70	83	86	12	251
64	41	49	64	6	173

At age 60, ρ_{x-1} is taken as $\rho_{59}+\rho_{60}^{(1)}$ and at age 61, as $\rho_{60}^{(2)}$. It should be noted that $E_{64}+\frac{1}{2}b_{64}-\frac{1}{2}e_{64}-\rho_{64}-\theta_{64}=123$, which gives correctly the number of retirements at exact age 65.

We next require E_x^d in order to compare the actual with the expected deaths by the A 1924–29 table, for the rates of mortality in the latter table are of course independent rates. This we obtain from the relation $E_x^d=E_x-\frac{1}{2}\rho_x$, excluding $\rho_{60}^{(1)}$ at age 60. The comparison is shown in the next table:

Age x	E_x (1)	E_x^d (2)	q_x A 1924/29 (3)	(2)×(3) (4)	θ_x (5)	q_x^d (6)
58	2500	2498	·01608	40·2	46	—
59	2355	2353	·01783	42·0	43	·01984
60	552	541	·01973	10·7	11	·02196
61	504	444	·02176	9·7	10	·02422
62	362	319	·02394	7·6	12	—
63	251	219	·02631	5·8	6	—
64	173	152	·02893	4·4	6	—
				120·4	134	

The percentage we require is therefore $\frac{134}{120{\cdot}4}\times 100=111{\cdot}3$ and column (6) gives the resulting rates of mortality for the active section of the service table, i.e. $1{\cdot}113\times q_x$ A 1924/29.

The next step is to calculate $q_x^{r(1)}$ and $q_x^{r(2)}$, the independent rates of retirement, the former being zero except at age 60. The relation $E_x^{r(2)}=E_x-\frac{1}{2}\theta_x$ gives the exposed to risk for $q_x^{r(2)}$, but $E_{60}^{r(1)}$ differs from the ordinary conception of exposed to risk in that it is the number of lives passing through exact age 60. The beginners at age 60 last birth-

day should not therefore contribute to $E_{60}^{r(1)}$, but the enders at age 60 should each contribute a whole unit. Hence

$$E_{60}^{r(1)} = E_{59}^{r(2)} + \tfrac{1}{2}b_{59} - \tfrac{1}{2}e_{59} - \rho_{59} - \tfrac{1}{2}\theta_{59}$$

and
$$E_{60}^{r(2)} = E_{60}^{r(1)} + \tfrac{1}{2}b_{60} - \tfrac{1}{2}e_{60} - \rho_{60}^{(1)} - \tfrac{1}{2}\theta_{60}.$$

The last value can, however, be obtained more simply from the general relation

$$E_{60}^{r(2)} = E_{60} - \tfrac{1}{2}\theta_{60}.$$

It is unnecessary to reproduce the working table for the exposed to risk. The values of this function and of the resulting rates are as follows:

Age x	$E_x^{r(1)}$	$\rho_x^{(1)}$	$q_x^{r(1)}$	$E_x^{r(2)}$	$\rho_x^{(2)}$	$q_x^{r(2)}$
59	—	—	—	2333·5	4	·00171
60	2248	1686	·7500	546·5	22	·04025
61	—	—	—	499	120	·24048

The next stage is to calculate the dependent rates so as to facilitate the construction of the service table. No account need be taken of $q_{60}^{r(1)}$ which does not affect the relation between the dependent and independent rates.

Age x	$1 - \tfrac{1}{2}q_x^{r(2)}$ (1)	$(1) \times q_x^d = aq_x^d$ (2)	$1 - \tfrac{1}{2}q_x^d$ (3)	$(3) \times q_x^{r(2)} = aq_x^{r(2)}$ (4)
59	·9991	·0198	·9901	·0017
60	·9799	·0215	·9890	·0398
61	·8798	·0213	·9879	·2376

We are now in a position to construct the service table by means of the rates in columns (2) and (4). The retirements at exact age 60 are obtained by multiplying $al_{60}^{(1)}$ by $q_{60}^{r(1)}$ and these retirements must be deducted from $al_{60}^{(1)}$ to obtain $al_{60}^{(2)}$. rl_x and rd_x represent survivors and deaths in the section of the table relating to lives who have retired.

Age x	$al_x^{(1)}$	$ar_x^{(1)}$	$al_x^{(2)}$	ad_x	$ar_x^{(2)}$	rl_x	rd_x
59	10,000	0	10,000	198	17	0	0
60	9,785	7339	2,446	53	97	7356	154
61	2,296	0	2,296	49	546	7299	170
62	1,701					7675	

The last two columns are obtained by the relations

$$rd_{59} = \tfrac{1}{2} ar_{59}^{(2)} \times q'_{59},$$

$$rl_{60} = ar_{59}^{(2)} + ar_{60}^{(1)} - rd_{59},$$

$$rd_{60} = (rl_{60} + \tfrac{1}{2} ar_{60}^{(2)}) \times q'_{60},$$

$$rl_{61} = rl_{60} + ar_{60}^{(2)} - rd_{60}, \text{ etc.,}$$

in which q'_{59}, etc., are the rates of mortality by the $a^{(m)}$ table.

The required probability is $\dfrac{rl_{62}}{al_{59}^{(1)}} = \cdot 7675.$

CHAPTER XX

CONSTRUCTION OF NATIONAL LIFE TABLES

1. The investigations that we have considered in previous chapters have all been based on data extracted from private records. We now turn to the problem of constructing mortality tables from national or public records.

National censuses have been taken in this country every 10 years since 1841 with the exception of 1941, when the census was postponed on account of the war. Compulsory registration of births, marriages and deaths was introduced in England in 1837 and in Scotland in 1854. Hence the data necessary for the construction of mortality tables from national statistics have been available for a century. The results of the English censuses have been combined in various ways with the registered deaths to construct the series of mortality tables known as English Life Tables* Nos. 1–10. The first three tables were constructed by Dr Farr, who did valuable pioneer work on this subject; both the theory and practice were greatly developed by George King, who constructed E.L.T. Nos. 7 and 8.

The information available from the censuses has gradually been increased. Prior to 1911 the population was recorded in quinquennial age groups, but from 1911 onwards data for individual ages are available. The data are also subdivided according to sex and, in the case of recent censuses, the females have been further divided by marital status. Records of births give the total number in each quarter subdivided according to sex, while the number of deaths is available for each calendar year, grouped according to age last birthday at death, sex and, in the case of females, marital status. In the first year of life, deaths are subdivided according to quarter years of age.

* Throughout this Chapter the abbreviation E.L.T. will be used for English Life Table(s).

2. Method of constructing tables.

It is plain that the data are well suited to the application of the census method of calculating mortality rates, and this is in fact the method which has been adopted. Before discussing details of the application of the method, it may be as well to consider why the more exact methods are unsuitable.

If we fix the limits of the experience as the dates of two adjacent censuses, the numbers of beginners and enders grouped according to age last birthday will be readily available. The deaths, grouped according to age last birthday at death, will also be given, while the main source of new entrants will be the births. In view of the way in which the deaths are grouped, an exact age (or life year) method is preferable to one based on age grouping by time (calendar year method), and the appropriate exposed to risk formulae are:

$$E_x = E_{x-1} + \tfrac{1}{2}(b_{x-1} + b_x) - \tfrac{1}{2}(e_{x-1} + e_x) - \theta_{x-1},$$

$$E_0 = \tfrac{1}{2}b_0 - \tfrac{1}{2}e_0 + n_0,$$

assuming an even distribution of birthdays over the calendar year for the lives included in b_x and e_x.

There are, however, other causes of movement which we have not taken into account, i.e. immigration and emigration. A certain amount of information about the number of migrants is recorded, but this does not include details of their ages. We should, therefore, have to make arbitrary assumptions about the ages of the migrants and, although the total number is not very great, the accuracy of the calculations would be seriously impaired.

The modified calendar year method (IX, 7) does not require any information about movements other than deaths between the annual dates on which the numbers under observation are recorded. This method is, however, inapplicable for two reasons: (*a*) censuses are taken only once every 10 years, whereas the method under discussion would need an annual census, and (*b*) deaths are not grouped according to their ages on the preceding census date, an essential feature of a calendar year method.

Let us now consider the application of the census method. Up to 1911, national life tables in this country were based on the deaths occurring in an intercensal period of 10 years and the corresponding

exposed to risk obtained from the two censuses. The central rate of mortality at exact age x was given by the expression $\dfrac{\theta_x}{\int_0^{10} {}^sP_x\,ds}$, where θ_x is the number of deaths at age x last birthday during the intercensal period and ${}^sP_x$ the population at age x last birthday at a point of time s years after the beginning of the period.

${}^0P_x$ and ${}^{10}P_x$ were the only known values of ${}^sP_x$, and these values were available for quinquennial age groups and not for individual ages. Two problems therefore arose:—

(*a*) It was necessary to make some more or less arbitrary assumptions as to the trend of the population. When dealing with so short a period as one year, the assumption that ${}^sP_x$ is a first degree function of s is not likely to introduce a serious error, but it is another matter to assume that it can be represented in this way over a period of 10 years.

(*b*) The mean population for quinquennial age groups and the corresponding deaths having been obtained, the rates of mortality for individual ages had to be calculated.

We shall now proceed to discuss these problems.

3. Estimates of the mean population.

Method 1. In constructing E.L.T. No. 6 based on the censuses of 1891 and 1901 and E.L.T. No. 7 based on those of 1901 and 1911, the assumption was made that the *total* population had increased in geometrical progression throughout the intercensal period, i.e. that at time s the population would be ${}^0Pr^s$ where $r^{10} = \dfrac{{}^{10}P}{{}^0P}$. This assumption was justified by Geo. King on the grounds that "population begets population". In more exact terms, the excess of the crude birth-rate (i.e. the ratio of the number of births in a calendar year to the total mean population during the year) over the crude death-rate was assumed to be constant throughout the intercensal period.

The assumption is not, of course, universally correct, and although in general it will be nearer the truth than the assumption

of an increase in arithmetical progression, it should be tested roughly by an examination of the number of births and deaths in each year covered by the investigation. Such a test was no doubt made before adopting the assumption when constructing E.L.T. Nos. 6 and 7.

The total mean population is given by one-tenth of the integral $\int_0^{10} {}^sP ds$ and, with the above assumption as to the rate of increase in the total population, this integral is easily reducible as follows:

$$\begin{aligned}\int_0^{10} {}^sP ds &= {}^0P \int_0^{10} r^s ds \\ &= \frac{{}^0P\,(r^{10}-1)}{\log r} \\ &= \frac{{}^{10}P - {}^0P}{\log r}.\end{aligned}$$

In order to obtain the mean population of the quinquennial age groups, the first inclination would be to make the same assumption as to the form of ${}^sP_{\bar{x}}$ (i.e. the population between exact ages $x-2$ and $x+3$) as was made for ${}^sP$. Unfortunately, the resulting mean populations do not add up to the total mean population, and this method is not therefore suitable. In constructing E.L.T. Nos. 6 and 7, use was made of a method devised by A. C. Waters, which assumes that the total population increases in geometric progression but that the ratio $\dfrac{{}^sP_{\bar{x}}}{{}^sP}$ is a first degree function of s.

Let $\alpha = \dfrac{{}^0P_{\bar{x}}}{{}^0P}$ and $\beta = \dfrac{{}^{10}P_{\bar{x}}}{{}^{10}P}$, whence $\dfrac{{}^sP_{\bar{x}}}{{}^sP}$ takes the form

$$\alpha + \frac{s}{10}(\beta - \alpha) = \alpha\left(1 - \frac{s}{10}\right) + \frac{\beta s}{10},$$

so that

$$\begin{aligned}{}^sP_{\bar{x}} &= {}^sP\left\{\frac{{}^0P_{\bar{x}}}{{}^0P}\left(1 - \frac{s}{10}\right) + \frac{{}^{10}P_{\bar{x}}}{{}^{10}P} \times \frac{s}{10}\right\} \\ &= \left(1 - \frac{s}{10}\right) r^s \times {}^0P_{\bar{x}} + \frac{s}{10} r^{s-10} \times {}^{10}P_{\bar{x}}.\end{aligned}$$

The central exposed to risk for the age group is therefore

$$\int_0^{10} {}^sP_{\bar{x}}\,ds = \int_0^{10} r^s\,{}^0P_{\bar{x}}\,ds - \int_0^{10} \frac{sr^s}{10}\,{}^0P_{\bar{x}}\,ds + \int_0^{10} \frac{sr^{s-10}}{10}\,{}^{10}P_{\bar{x}}\,ds.$$

The last two integrals may be evaluated by the process of integration by parts and the result is as follows:

$$\frac{(r^{10}-1)}{\log r}\,{}^0P_{\bar{x}} - \left(\frac{r^{10}}{\log r} - \frac{r^{10}-1}{10\,(\log r)^2}\right){}^0P_{\bar{x}} + \left(\frac{1}{\log r} - \frac{1-\dfrac{1}{r^{10}}}{10\,(\log r)^2}\right){}^{10}P_{\bar{x}}$$

$$= \frac{{}^0P_{\bar{x}}}{\log r}\left(\frac{r^{10}-1}{10\log r} - 1\right) + \frac{{}^{10}P_{\bar{x}}}{r^{10}\log r}\left(r^{10} - \frac{r^{10}-1}{10\log r}\right).$$

This function is in the form $10l\,{}^0P_{\bar{x}} + 10m\,{}^{10}P_{\bar{x}}$, where l and m are independent of x. It is, therefore, a simple matter to evaluate the mean population $l\,{}^0P_{\bar{x}} + m\,{}^{10}P_{\bar{x}}$ for any age group once l and m have been obtained.

Method 2. The above method is an arbitrary one for which there is no theoretical justification, but in general the error involved at any age group will be unimportant, provided that the basic assumption regarding the rate of increase in the total population is roughly correct. It has, however, one possible defect. If ${}^0P_{\bar{x}} = {}^{10}P_{\bar{x}}$, the obvious course is to assign this value to the mean population unless evidence to the contrary is forthcoming, but Waters's original method would produce a different value, unless $r=1$.

To overcome this defect, Waters devised another method whereby l and m were obtained by the equations

$${}^0Pl + {}^{10}Pm = \bar{P}, \quad l+m=1,$$

$\bar{P}$ being the mean population for all ages combined. The mean population for age group $x-2$ to $x+3$ is then taken to be

$$l\,{}^0P_{\bar{x}} + m\,{}^{10}P_{\bar{x}}.$$

Waters states that this method gives results which differ very little from those according to his earlier method, and the later method is therefore preferable on account of its simplicity.

Method 3. In constructing E.L.T. Nos. 8, 9 and 10, the problem was largely avoided. The deaths occurring in a period of only three years instead of 10 years were used, the period being chosen so that

its middle point would fall near a census date. The population according to this census was employed with little or no modification as the mean population over the three-year period. This process assumes that ${}^{s}P_{\bar{x}}$ can be expressed as a first degree function of s, an assumption which is in general sufficiently near the truth over a period as short as three years.

E.L.T. No. 8 was based on the 1911 census and the deaths in 1910–12. The census was taken in April and adjustments were made by reference to the 1901 census population to allow for the fact that the census date did not coincide with the mid-point of the three years. In 1921 the census was taken in June and it was not thought necessary to make any adjustments, nor were any adjustments made in 1931 although the census was taken in April.

4. Rates of mortality at individual ages.

The accepted method of obtaining rates of mortality for individual ages from the mean populations and the corresponding deaths for quinquennial age groups was devised by Geo. King and described in Part I of the supplement to the *75th Annual Report of the Registrar-General.* It involved the calculation of the numbers of the population and deaths applicable to individual quinquennial ages. Hence he obtained for those ages the values of m_x, which he called quinquennial pivotal values, and interpolated for the remaining values of m_x. A full description of these processes is outside the scope of this book, as questions of graduation are involved, and it is sufficient to say that these methods, first applied in constructing E.L.T. Nos. 7 and 8, were subsequently used by other operators for E.L.T. Nos. 9 and 10. Special processes were necessary at certain ages and these will be dealt with later in the chapter.

It is worth noting that, in constructing E.L.T. No. 7, no use was made of the population data for individual ages recorded at the 1911 census, seeing that similar information was not forthcoming from the 1901 census.

For E.L.T. No. 8 and later tables, figures for individual ages were available instead of for quinquennial age groups only. Examination of these data showed that there appeared to be a tendency for members of the public to favour certain figures,

i.e. a figure ending in 0 or 5 and to a less extent an even number in preference to an odd. This feature is more prevalent in the census than in the death returns and, if the unadjusted data had been used to calculate the rates of mortality, errors would have been introduced. Attempts were accordingly made to eliminate the effect of these misstatements of age by combining the data in quinquennial age groups in order to achieve a balance of errors. In the case of E.L.T. No. 10, for example, the most popular endings in the population data were 0 and 8 while 1, 7 and 9 were the least popular, and the age groups adopted were those ending in the digits 0–4 and 5–9. For E.L.T. Nos. 8 and 9, the corresponding groups were those ending in the digits 4–8, 9–3 and 2–6, 7–1.

The data having been grouped in this way, quinquennial pivotal values were calculated as before and hence values of m_x were obtained.

5. Errors in census data.

There are a number of sources of error in census and death returns in addition to the misstatements of age mentioned in the previous paragraph.

(*a*) There is a tendency for women, especially those in middle life, to understate their ages deliberately in the census returns and, as these misstatements are not likely to be present to the same extent in the death returns, the rates of mortality are underestimated at the ages where the exposed to risk is overstated and overestimated at the older ages where it is correspondingly understated. It is obviously impossible to correct this error or to make more than a very rough estimate of its effect.

(*b*) At the old ages, there is a tendency for both men and women to overstate their ages in the census returns, with the result that the rates of mortality at the very old ages are underestimated and those at the youngest ages where the feature exists overestimated. It is difficult to adjust for the error at the latter ages, but at the very old ages, where the rates are somewhat arbitrary, they can be adjusted in the process of graduation if the rate of increase is obviously too slow. It has been stated that, with improved educa-

tion, the feature has not been present to the same extent in the last two censuses, but this is largely a matter of conjecture.

(*c*) At the young ages, comparison of the census figures with the population figures built up from the returns of births and deaths has revealed an unaccountable deficit in the former. This feature of the data first attracted attention in E.L.T. No. 6, but it was evidently assumed that there must be some satisfactory explanation. The census total for ages 0–4 inclusive was accordingly used without adjustment, although the numbers were redistributed among the individual ages in proportion to the population figures built up from the returns of births and deaths. For consistency, Geo. King followed the same procedure in E.L.T. No. 7, although he considered that the deficit must have been due to the omission of a considerable number of children under age 2 from the census returns. In E.L.T. No. 8, however, he neglected the census figures altogether and derived the exposed to risk directly from the returns of births and deaths. The same course has been followed in the construction of E.L.T. Nos. 9 and 10. The exact process deserves careful consideration.

6. Exposed to risk at young ages.

Let β^k be the number of births and θ_0^k the number of deaths at age 0 last birthday in the calendar year k. Then, by the method of Chapter II, the rate of mortality at age 0 in the years $k-1$, k and $k+1$ is

$$q_0=\frac{\theta_0^{k-1}+\theta_0^k+\theta_0^{k+1}}{(1-s^{k-2})\,\beta^{k-2}+\beta^{k-1}+\beta^k+s^{k+1}\beta^{k+1}}, \qquad \ldots\ldots(1)$$

where s^k is the average period between the dates of birth in the year k and the end of that year. The assumption is made that the effect of migration is negligible and this is probably justified at the very young ages.

For age 1, the corresponding expression is

$$q_1=\frac{\theta_1^{k-1}+\theta_1^k+\theta_1^{k+1}}{(1-s^{k-3})\,\beta^{k-3}+\beta^{k-2}+\beta^{k-1}+s^k\beta^k-(\text{a term in }\theta)}. \qquad \ldots\ldots(2)$$

The last term will be based on deaths before the attainment of

age 1 among the lives included in the β terms of the denominator. $1-s^{k-3}$ must be deducted for each death among β^{k-3}, s^k for each death among β^k, and a unit for each other death.

Unfortunately, there is no record of the deaths occurring among the lives born in a particular calendar year and an approximation is necessary. Some of the deaths at age 0 among β^{k-3} will occur in year $k-3$, the remainder in $k-2$ and similarly for the other years of birth. The majority of the deaths with which we are concerned will therefore relate to the years $k-2$, $k-1$ and k, and we might approximate to the required term in θ by taking the total of the deaths at age 0 in these three years. As the term in θ will be small relative to the total denominator, the approximation should be sufficiently close unless either s^{k-3} and s^k or β^{k-3} and β^k differ from each other very considerably.

For age 2, the denominator would be based on the births in the years $k-4$ to $k-1$, and the approximate deduction would consist of the number of deaths at age 0 in the years $k-3$, $k-2$ and $k-1$ and the deaths at age 1 in the years $k-2$, $k-1$ and k. The denominators for ages 3 and 4 would be similar.

In constructing E.L.T. No. 8, King assumed an even distribution of births over the calendar year so that s^{1905}, s^{1906} etc. were each taken to be $\frac{1}{2}$. The approximation regarding the deaths referred to above was also made and was justified by the fact that the birth-rate did not vary much over the years 1905–12. The rate of mortality at age 2, for example, was taken to be

$$\frac{\theta_2^{1910}+\theta_2^{1911}+\theta_2^{1912}}{\frac{1}{2}\beta^{1907}+\beta^{1908}+\beta^{1909}+\frac{1}{2}\beta^{1910}-(\text{deaths at age 0 in 1908–10} + \text{deaths at age 1 in 1909–11})} \qquad \ldots\ldots(3)$$

In 1921, a new factor had to be taken into account, i.e. the violent fluctuations in the birth-rate in the years during and immediately following the war of 1914–18. The assumption that the births were evenly spread over each calendar year would have introduced quite substantial errors (X, 4). Account was therefore taken of the returns of births for the appropriate quarters, the assumption being made that the births in each quarter were

uniformly spread. The formula for the rate of mortality at age 0 was

$$q_0 = \frac{\theta_0^{1920} + \theta_0^{1921} + \theta_0^{1922}}{\begin{array}{r}\frac{1}{8}\left({}^1\beta^{1919} + 3\,{}^2\beta^{1919} + 5\,{}^3\beta^{1919} + 7\,{}^4\beta^{1919}\right) + \beta^{1920} + \beta^{1921} \\ + \frac{1}{8}\left(7\,{}^1\beta^{1922} + 5\,{}^2\beta^{1922} + 3\,{}^3\beta^{1922} + {}^4\beta^{1922}\right)\end{array}}, \quad \ldots\ldots(4)$$

where ${}^t\beta^k$ is the number of births in the tth quarter of the year k. For age 2, the rate would be

$$q_2 = \frac{\theta_2^{1920} + \theta_2^{1921} + \theta_2^{1922}}{\begin{array}{r}\frac{1}{8}\left({}^1\beta^{1917} + 3\,{}^2\beta^{1917} + 5\,{}^3\beta^{1917} + 7\,{}^4\beta^{1917}\right) + \beta^{1918} + \beta^{1919} + \frac{1}{8}\left(7\,{}^1\beta^{1920}\right. \\ \left. + 5\,{}^2\beta^{1920} + 3\,{}^3\beta^{1920} + {}^4\beta^{1920}\right) - (\text{deaths at age 0 in 1918–20} \\ + \text{deaths at age 1 in 1919–21})\end{array}}. \quad \ldots\ldots(5)$$

A similar process was carried out in constructing E.L.T. No. 10. Here a refinement was introduced at age 0 to reduce the error arising from the impact of a force of mortality which falls very rapidly between exact ages 0 and 1 on an exposed to risk which, owing to differences in the birth-rates in 1929 and 1932, would not be distributed over the year of age in the normal pattern (see VIII, 2). The methods used in the last few paragraphs assume an even distribution of deaths over the year of age, and the modified method assumed an even distribution over each quarter year of age in the first year of life.

For this purpose, the probabilities of a child aged 0 exactly dying in the first, second, third and fourth three months of life were calculated and the sum of the four probabilities gave the value of q_0. The formulae were

$$q_0^{(0-3\text{ months})} = \frac{\text{Deaths in 1930–32 (age 0–3 months)}}{\frac{1}{2}\,{}^4\beta^{1929} + \beta^{1930} + \beta^{1931} + \beta^{1932} - \frac{1}{2}\,{}^4\beta^{1932}}, \quad \ldots\ldots(6)$$

$$q_0^{(3-6\text{ months})} = \frac{\text{Deaths in 1930–32 (age 3–6 months)}}{\frac{1}{2}\,{}^3\beta^{1929} + {}^4\beta^{1929} + \beta^{1930} + \beta^{1931} + \beta^{1932} - \frac{1}{2}\,{}^3\beta^{1932} - {}^4\beta^{1932}},$$

and so on. ......(7)

7. Possible sources of error.

Besides the errors referred to in paragraph 5 which are inherent in the census data, errors may arise due to certain other factors of which our methods may not take full account. Most of these, as

we shall see, affect the exposed to risk principally at the younger ages:

(*a*) The births are not normally spread evenly over the calendar year.

(*b*) The force of mortality in the first few months of life varies slightly over the calendar year.

(*c*) The force of mortality in the first few years of life, especially in the first few months, falls rapidly as age advances.

(*d*) The returns of births and deaths in a given year or quarter year are based on the numbers registered which, through delays in registration, may not correspond to the numbers actually born or dying in the period.

Let us consider exactly how these factors give rise to errors and how far any corrections made in the past should be adequate.

(*a*) *Distribution of births over calendar year.* As we have already seen, the assumption of a uniform distribution was made at the young ages in E.L.T. No. 8 (see the terms $\frac{1}{2}\beta^{1907}$ and $\frac{1}{2}\beta^{1910}$ in formula (3)). Even if the distribution were not a uniform one, the error involved would be negligible provided that the number and distribution of the births were roughly the same in the two years in question.

In formula (4) the assumption is made that over each quarter in the years 1919 and 1922 the distribution of births was uniform. Obviously this modification should reduce the errors substantially.

The influence of this factor is not, however, limited to the young ages. When the data consist of a single census and three years' deaths, the assumption underlying the census method is that ${}^sP_x$ is a first degree function of s over the three years. In the light of the discussion in (X, 4), we arrived at the conclusion that this assumption is sufficiently correct in the case of national statistics, unless the birth-rate has undergone considerable fluctuations in the years in which the lives under review were born.

As an example of a period during which the birth-rate fluctuated violently, we shall once again take the years during and shortly after the war of 1914–18. The fluctuations affected the population up to age 6 at the 1921 census, but a census method was unsuitable

at these ages for other reasons. At the 1931 census, the population aged from 10 to 16 was affected and the ordinary census method would normally have been applied at these ages. One possible solution of the problem was to use the method employed at the very young ages. This would not, however, have been wholly satisfactory for, except in the first few years of life, it is unsound to neglect the effect of migration.

An alternative solution would be to take as the central exposed to risk at age x the sum of the census populations at ages $x-1$, x and $x+1$. The rationale of this method is that ${}^{s}P_{x+1}$ is an approximation to ${}^{s-1}P_x$ and ${}^{s}P_{x-1}$ an approximation to ${}^{s+1}P_x$, and that the errors in these two approximations are of opposite sign. Hence, the sum of the three populations should be a good approximation to the sum of the populations at age x on the census date in 1931 and on its anniversaries in 1930 and 1932, and the latter sum should provide a close estimate of the central exposed to risk over the three years. The 1931 census took place about one-third of the way through the year, so that the latest method can be expressed in symbols as follows:

$$\begin{aligned}\int_0^3 {}^{s}P_x\,ds &\fallingdotseq {}^{\frac{1}{3}}P_x + {}^{\frac{4}{3}}P_x + {}^{\frac{7}{3}}P_x \\ &\fallingdotseq {}^{\frac{4}{3}}P_{x+1} + {}^{\frac{4}{3}}P_x + {}^{\frac{4}{3}}P_{x-1}. \qquad \ldots\ldots(8)\end{aligned}$$

This process takes into account variations in the number of births between one year and another, but does not allow for variations in the birth-rate over the calendar year.

The Government Actuary overcame the latter objection by means of an ingenious adjustment. The adjustment for age 12, for example, is as follows.

The lives included in expression (8) are the survivors of the births in the years 1918 and 1919 and some of the births in 1917 and 1920. On the assumption of an even distribution in the first and last quarters the total number of these births is:

$$\tfrac{2}{3}\,{}^{2}\beta^{1917} + {}^{3}\beta^{1917} + {}^{4}\beta^{1917} + \beta^{1918} + \beta^{1919} + {}^{1}\beta^{1920} + \tfrac{1}{3}\,{}^{2}\beta^{1920} = B_{12}.$$

The following expression, however, includes each group of births to the extent that the survivors of the group contribute to the correct exposed to risk corresponding to the deaths in 1930–32 at

age 12, assuming an even distribution of births in each quarter of 1917 and 1920:

$$\tfrac{1}{8}({}^{1}\beta^{1917}+3\,{}^{2}\beta^{1917}+5\,{}^{3}\beta^{1917}+7\,{}^{4}\beta^{1917})+\beta^{1918}+\beta^{1919}$$
$$+\tfrac{1}{8}(7\,{}^{1}\beta^{1920}+5\,{}^{2}\beta^{1920}+3\,{}^{3}\beta^{1920}+{}^{4}\beta^{1920})=A_{12}.$$

The exposed to risk from formula (8) was accordingly adjusted by multiplying by $\dfrac{A_{12}}{B_{12}}$, so that the value of m_{12} was

$$\frac{\text{Deaths at age 12 in 1930–32}}{\text{Census population at ages 11, 12 and 13}}\times\frac{B_{12}}{A_{12}}.$$

This method was used at ages 6–16, and smooth junctions effected at the lower limit with the rates for the very young ages calculated by the life-year method and at the upper limit with the rates calculated by the ordinary census method.

The complications introduced by the fluctuating birth-rate during and after the war of 1914–18 will be met at ages 10 years older at each successive decennial census. It is doubtful whether the method employed in 1931 will be equally suitable as the ages involved become older and it will be interesting to see how the problem is solved.

(*b*) *Variation of μ_x over calendar year.* At all ages the rate of mortality tends to be higher in the winter months, and this applies in particular to infantile mortality. As the force of mortality in the first few weeks after birth is comparatively high, it is clear that variations in the distribution of the births over the calendar year will cause variations in the rate of mortality at age 0. In general, however, the inclusion of three calendar years in the experience should reduce considerably the effect of these variations in the distribution of births.

(*c*) *Variation of μ_x over the year of age.* As indicated in paragraph 6, some account was taken of this factor in 1931, but the method used did not make any allowance for the very rapid fall in the force of mortality between ages 0 and 3 months. The error in formula (6) from this cause would be negligible if the number and distribution of births were roughly the same in the last quarters of 1929 and 1932.

The problem of the first two years of life was dealt with more

thoroughly by the Registrar-General in his 1934 Statistical Review. His method was to estimate the number of births corresponding to the tabulated number of deaths at each age in months, from which he obtained the necessary probabilities to build up a life table showing the survivors of a given number of births at the end of each month up to the twenty-fourth. Hence, he estimated that, of the deaths at ages 0–3 months registered in a particular quarter, 1/12th would be among children whose births were registered in the preceding quarter and the remainder would be among children whose births were registered in the quarter in which death occurred.

Applying this result to modify formula (6), we get as our denominator

$$\tfrac{1}{12}\,{}^{4}\beta^{1929}+\beta^{1930}+\beta^{1931}+\beta^{1932}-\tfrac{1}{12}\,{}^{4}\beta^{1932}. \qquad \ldots\ldots(9)$$

Obviously, if ${}^{4}\beta^{1929}={}^{4}\beta^{1932}$, the two formulae give identical results. A similar process can be followed for the probability of death at age 3–6 months, but the error in formula (7) is much less serious than that in formula (6).

(*d*) *Delay in registration.* The returns of births for a given quarter are based on the number registered in that quarter and, as a period of a few weeks on the average elapses between birth and registration, some of these births will have taken place in the previous quarter. As the corresponding time lag for deaths is only about two days, the denominators of the fractions in the foregoing formulae are not entirely based on the births from which the deaths in the numerators arose. The Registrar-General made some allowance for this in the estimate described in (*c*) above, but it is not clear how his conclusions were reached. He points out that, of 100 children dying aged 0–3 months in the July–September quarter, the distribution of births will have been approximately as follows:

Month	Number of Births
April	2
May	6
June	15
July	32
August	27
September	18
	100

He then argues that, allowing for a month's delay in the registration of the corresponding births, only those births occurring in April and May would be registered in the second quarter of the year, i.e. about 1/12th of the total births out of which third quarter deaths at ages 0–3 months could arise. He does not, however, appear to take account of the fact that the September births, amounting to nearly 1/5th of the total, would not be registered until October on the average. Hence formula (9) would presumably be more accurate if the term $\frac{1}{5}\,{}^{1}\beta^{1933}$ were added and the term $\frac{1}{5}\,{}^{1}\beta^{1930}$ deducted. The error is not serious and it is difficult to obtain a satisfactory adjustment, so long as births and deaths are tabulated according to date of registration instead of date of birth or death. The Government Actuary was no doubt referring to difficulties of this kind when he stated in Part I of the *Decennial Supplement*, 1931, that any attempt to obtain an accurate exposed to risk would necessitate assumptions for which no authority could be claimed.

8. Population statistics: some general problems.

We shall conclude the chapter by considering some general problems concerning mortality rates and population statistics.

(*a*) *Estimates of population between censuses.* We have already met the problem of estimating the population at each age at a particular date between censuses. The solutions put forward so far have, however, assumed that ${}^{s}P_x$ can be expressed as a simple function of s, and circumstances may arise to make this assumption untenable. Let us suppose that we wish to estimate the population half-way between two decennial censuses.

One method would be to obtain ${}^{5}P_{x+5}$ from ${}^{0}P_x$ by deducting the appropriate deaths. If we assume that the census was taken four months after the beginning of the calendar year k, we require first of all the deaths among ${}^{0}P_x$ in the remaining two-thirds of the year, some of which would occur at age x and some at age $x+1$ last birthday. We cannot trace the deaths which actually relate to ${}^{0}P_x$, but the expression $\frac{2}{3}\left(\frac{2}{3}\theta_x^k+\frac{1}{3}\theta_{x+1}^k\right)=\frac{4}{9}\theta_x^k+\frac{2}{9}\theta_{x+1}^k$ is a fair approximation, for all the deaths among ${}^{0}P_x$ occurring on the census date are

included in θ_x^k and, of the deaths occurring at the end of the year k, about two-thirds are included in θ_{x+1}^k and one-third in θ_x^k, so that on the average two-thirds of the deaths among ${}^0P_x$ between the census date and the end of the year are included in θ_x^k and one-third in θ_{x+1}^k. In the year $k+1$, the deaths to be deducted would, by the same process, be $\frac{1}{3}\left(\frac{1}{6}\theta_x^{k+1}+\frac{5}{6}\theta_{x+1}^{k+1}\right)$ before the anniversary of the census date and $\frac{2}{3}\left(\frac{2}{3}\theta_{x+1}^{k+1}+\frac{1}{3}\theta_{x+2}^{k+1}\right)$ thereafter, making a total of $\frac{1}{18}\theta_x^{k+1}+\frac{13}{18}\theta_{x+1}^{k+1}+\frac{2}{9}\theta_{x+2}^{k+1}$. The deductions would be similar for the years $k+2$, $k+3$ and $k+4$, while in the year $k+5$ the deduction would be $\frac{1}{18}\theta_{x+4}^{k+5}+\frac{5}{18}\theta_{x+5}^{k+5}$. The first approximation to ${}^5P_{x+5}$ would therefore be

$$\begin{aligned}{}^0P_x-\tfrac{1}{18}\,(8\theta_x^k+4\theta_{x+1}^k+\theta_x^{k+1}+13_{x+1}^{k+1}+4\theta_{x+2}^{k+1}+\theta_{x+1}^{k+2}+\ldots\\ +4\theta_{x+5}^{k+4}+\theta_{x+4}^{k+5}+5\theta_{x+5}^{k+5}).\end{aligned}$$

A second approximation could be made by adding to ${}^{10}P_{x+10}$ the appropriate deaths between ages $x+5$ and $x+10$, and the reader should have no difficulty in deriving the following expression:

$${}^{10}P_{x+10}+\tfrac{1}{18}\,(8\theta_{x+5}^{k+5}+4\theta_{x+6}^{k+5}+\theta_{x+5}^{k+6}+\ldots+5\theta_{x+10}^{k+10}).$$

These two expressions are unlikely to be identical, for the methods are only approximate and no allowance has been made for migration. The mean of the two should, however, be a fairly good estimate of ${}^5P_{x+5}$.

Another method employs probabilities of survival instead of registered deaths. These probabilities would have to be estimated from the corresponding probabilities according to the life tables constructed from the census data. We might assume that the mortality rates at a given age had changed in arithmetic progression between the two censuses, and it would then be a simple matter to obtain the probabilities for each year and so build up the values of ${}_5p_x$ and ${}_5p_{x+5}$ necessary to make the two approximations to ${}^5P_{x+5}$. The final estimate would once again be the mean of the two approximations.

It must be remembered that we have assumed that ${}^sP_x$ cannot be expressed as a simple function of s, and this probably means that there has been some disturbing factor during the decennium. Hence the assumption that the rates of mortality for a given age have changed in arithmetic progression during the decennium may well be incorrect. In that event, unless data are available to

permit of a more accurate assumption being formulated, the method based on the use of deaths should give more reliable results than that based on probabilities of survival.

Whichever method be employed, recourse must be had to the registered births in order to obtain an approximation to the population at very young ages.

(*b*) *Estimates of future population.* In view of the falling birth-rate a problem which has received a good deal of attention in recent years is the estimation of the size and age distribution of the population in the future. For this purpose we must make some assumptions as to the trend of both birth and mortality rates. Obviously any such assumptions, particularly as regards birth-rates, must be largely a matter of conjecture and, instead of making a single estimate of the population, it is desirable to make at least two estimates which will give a fairly safe indication of the upper and lower limits within which the population may be expected to lie.

In trying to decide the trend of future birth-rates, it would be unsound to work on the crude birth-rate, for this rate is a compound function depending on two variables, i.e. the proportion of women of child-bearing age in the population and the birth-rate among these women. A fall in the female birth-rate in 1930, for example, has therefore a dual effect on the estimates—it causes a direct fall in the numbers at age 10 in 1940, age 20 in 1950, etc., compared with the corresponding numbers at these ages in the preceding years, and from 1950 to 1975 or thereabouts it causes a further reduction by reducing the number of potential mothers.

A suitable method of procedure is therefore to fix our assumptions as to the trend of the birth-rate, taken as the ratio of the number of births to the number of women aged between 20 and 45, and then to apply the appropriate rate for each calendar year to the estimated number of women between these ages in that year. In tackling problems of this nature involving approximate estimates, it is of advantage to make two approximations representing the upper and lower limits between which the result may lie. With this object in view, we may assume

(*a*) a gradually falling birth-rate tending to a lower limit, and

(*b*) a gradually rising rate tending to an upper limit.

(*a*) will give a lower limit for the future population and (*b*) an upper limit.

The number of births would have to be divided according to sex, but as the proportion varies comparatively little from year to year this should not present any serious difficulties.

It may be suggested that it would be more accurate to use issue rates applicable to the age of the mother, in which case we should have to estimate the proportion of married women among the women at each age. Unfortunately, owing to deficiencies in our national statistics, reliable issue rates based on age are not yet available, although the additional information now obtained when a birth is registered will in time remedy this defect. In any case, it is doubtful whether a refinement of this kind would be justified when we require an estimate only on very broad lines.

Assumptions must also be made as to the trend of future mortality rates, a problem which has already been discussed in Chapter XVI. For our lower limit, we might assume that the rates according to the latest national life table would remain in force, but assume for the upper limit a continued improvement in mortality. In either case, we should probably advance by periods of 10 years, using group probabilities of the form ${}_{10}p_{7\frac{1}{2}}$, ${}_{10}p_{12\frac{1}{2}}$, etc., for age groups 5–9, 10–14, etc., as it would obviously be out of place to introduce greater refinements except at the very young ages where we should use factors of the form ${}_{10}p_{\frac{1}{2}}$, ${}_{10}p_{1\frac{1}{2}}$, etc.

It is hardly necessary to say that, in presenting the results of any such investigation, the nature of the assumptions should be clearly stated.

CHAPTER XXI

NATIONAL LIFE TABLES. SECTIONAL INVESTIGATIONS

1. Types of sectional investigation.

From 1911 onwards, the Decennial Supplement to the Registrar General's report has included the results of an investigation into the mortality of different classes of the population in addition to the tables applicable to the male and female populations as a whole. The factors whose effects have been examined are marital status, density of population, geographical situation, occupation and social class and in all cases males and females have been classified separately.

2. Marital status.

The subdivision of the data into marital status has been made for females only. Separate life tables for spinsters, married women and widows were constructed by King from the 1911 census and the deaths of 1910–12, but Watson did not follow this course in the two succeeding decennial investigations and published only the different rates of mortality.

Watson argued that a life table for spinsters is misleading, as it suggests that the mortality shown at age $x+t$ is that applicable to the survivors of the spinsters, who, t years earlier, were aged x. A proportion of these spinsters will have married in the meantime and, unless the mortality rates for spinsters and married women are identical—in which case there is no point in keeping the classes separate—the rate for spinsters at age $x+t$ will not in general apply to the survivors of those aged x. Similar considerations arise for married women, since the class is subject to increment by spinsters who marry and to decrement by the deaths of husbands, and also for widows, since the class is subject to increment by the deaths of husbands and to decrement by remarriages.

There is no gainsaying this argument, but it may be pointed out that a similar objection applies in many other ways. Compare, for

example, the operation of the time factor on a life office experience as a result of which the rates for age $x+t$ in 1921 are unlikely to be experienced in $1921+t$ by the survivors of the lives aged x in 1921. The argument may therefore be extended in criticism of the great majority of life tables, as we have already seen in Chapter XVI.

At first sight, it might be thought that the fact of mortality rates varying according to marital status makes it unsound on the grounds of heterogeneity to calculate rates for the female population as a whole. It is certainly true that this fact is reliable evidence of heterogeneity, provided of course that the differences are statistically significant, but to justify the criticism we must show that marital status has a direct influence on mortality in the sense that a change in the proportion of married women would affect the rates of mortality for the female population as a whole.

Now it is clear that the mere fact of a woman being married will have some influence on her longevity. On the one hand, she will in general enjoy a more natural and regular life than a spinster; on the other hand, she will be exposed to the risks of child-birth. Similarly, the heavier mortality experienced by widows compared with the rates for married women is no doubt largely due to the less favourable conditions to which they are normally subject. To this extent, the criticism is sound.

As against this, however, one of the chief factors influencing the relative mortality of single and married women is the selection exercised at marriage, whereby the healthier lives are more likely to marry. An increase in the number of marriages would tend to lower the standard of selection and would not, so far as this factor is concerned, affect the mortality experience of the whole female population. In fact, the existence of the selective influence is merely a result of the obvious fact that all lives at the same age do not enjoy the same state of health.

The criticism is therefore partly but not wholly justified. In any case, there are, as we shall soon see, other factors making for a far higher degree of heterogeneity. Obviously, the population of a country must from its very nature be a heterogeneous body.

3. Density of population.

It is a matter of common knowledge that country life is in general more healthy than life in the towns. King tried to obtain a measure of this influence by investigating separately the mortality of the county boroughs, urban districts and rural districts and of the county of London. For both sexes, the rural districts showed the lightest mortality in nearly all age groups, while the county boroughs or the county of London usually showed the heaviest mortality.

Similar divisions were adopted by Watson in his investigations ten and twenty years later.

4. Geographical district.

Subdivisions on these lines were first made by Watson in his investigation of the 1921 census data. He chose eleven districts which, in the light of his wide experience of such matters, seemed likely to provide fairly homogeneous data so far as geographical influences are concerned. The mortality of each district was investigated separately for county boroughs, urban districts and rural districts and, of course, for each sex separately, so that a considerable number of classes were segregated. The results showed a considerable degree of difference in the mortality rates experienced by the various districts.

It was not thought necessary to undertake the heavy work which the construction of a complete set of rates for each subdivision would have involved, but rates were calculated for the subdivisions showing the heaviest and the lightest mortality, i.e. Northumberland and Durham county boroughs and Eastern Counties rural districts. Complete life tables were constructed for Greater London.

A similar course was followed ten years later. In the meantime, the Registrar-General had divided the country into regions for purposes of his annual statistical review and these were accordingly used instead of the former districts.

5. Occupation.

Ever since the 1851 census, investigations have been made to test the effect of occupation, but until the 1921 census the classification was on an industrial rather than a true occupational basis. For

example, a clerk working for a colliery company was formerly included in the coal mining industry, but under the new method of classification he was included as a clerk.

Occupation has a greater influence on mortality than does the nature of the industry with which a man is associated. Hence the new method tended to produce a greater difference between the rates for the various classes, since the old method involved a certain amount of averaging due to the inclusion in one class of men engaged in different occupations subject to different rates of mortality. The advantage of a true occupational classification is obvious if we consider an industry involving a heavy extra risk which does not apply in the same degree to employees engaged in different occupations in the same industry. Cutlery grinders, for example, are subject to a special silica risk which does not apply to other workers in the cutlery industry.

6. Occupation: mortality indices.

We have already discussed the methods by which comparison of the rates experienced by different occupations has been simplified (Chapter XVIII). We have not, however, discussed fully the range of ages which such a comparison should cover, nor the factors which have influenced the choice of the standard population or the standard rates of mortality required to calculate mortality indices.

It is generally agreed that, for the purpose of calculating a mortality index which is intended to be representative of the occupation as a whole, the data should be limited to the ages at which men are actively engaged in employment. It is not suggested that occupation ceases to have any effect on mortality once a man has retired, as obviously his length of life after retirement will depend on his state of health when he retired, which in turn depends on the occupation that he followed. Nevertheless, the effect of occupation will not be present to the same degree and as age advances it will gradually disappear.

Prior to 1921, mortality indices were based on ages 25–65, but this range was extended by the inclusion of the age group 20–25 in the investigation based on the deaths of 1921–23. These years were used instead of 1920–22, as it was thought that in 1920 many

men who had only recently been demobilized would not have settled down to a fixed occupation. This meant that the census population, used as the central exposed to risk, applied to a date near the beginning of the three year period. The justification for the inclusion of the earlier age-group was that by age 20 young men have been subject to the influences of their occupations sufficiently long for these influences to affect their mortality.

The change had the advantage of increasing the volume of data. Further, what is more important, ages were included at which certain diseases are specially prevalent. As these diseases are more common in some occupations than in others, the change had some effect on the relative positions of the various occupations. On the other hand, the numbers for some occupations were small, the results were unreliable and the errors were magnified by the method used for calculating the C.M.F.

In the 1931 Decennial Supplement, it was pointed out that some of the objections to the inclusion of ages 20–25 also apply to the ages up to about 35. The ratio of actual to expected deaths was accordingly given for ages 35–65 as well as for ages 20–65. This ratio, as opposed to the C.M.F., was the principal index used in 1931; a comparison of the values of the two indices for the principal occupations is given in the report.

As regards the choice of a standard population and standard mortality rates, the experience of the whole male population between the limiting ages was used in 1911, but in 1921 only the occupied and retired civilian males were included. The difference was small, but it was thought better to exclude those who had never engaged in any occupation; in this class would be included those who had been invalid or insane since childhood.

The mortality rates experienced by all males and by the occupied and retired civilian males show a marked divergence above age 70, the former being much lighter than the latter. This phenomenon was attributed to the tendency for old men to omit their former occupations from the census schedules. As the Registrar's enquiries would usually elicit these particulars on a death being registered, the rates of mortality for the occupied and retired were overstated. This is a further reason for excluding the old ages when

calculating mortality indices. The whole question is discussed very fully in Appendix A of Part II of the Registrar-General's Decennial Supplement, 1921.

In the case of the 1931 census figures, the rates for the total male population aged 20–65 were taken as the standard. While one reason for this was that there was less disparity between the rates for all males and for all occupied and retired civilian males than was the case ten years earlier, the main reason was to achieve consistency with the simultaneous investigation into the mortality of married women according to the occupations of their husbands and of single women according to their own occupations. For the married women, it was desirable to include the wives of the unoccupied in the investigation; for single women, exclusion of the unoccupied would have reduced the standard data by about 20 per cent.

Incidentally, the rates of mortality at the old ages among the occupied and retired males were not nearly so much in excess of those for all males as was formerly the case. This suggests that there was more careful completion of the census schedules.

7. Occupation: errors.

Two types of error in occupational mortality investigations deserve attention. One is the tendency to magnify the importance of a man's occupation, e.g. to describe a clerk as a departmental head. This tendency would be less important if it were equally prevalent in census and death returns, but in fact the tendency is much stronger in the case of the former. Hence the rates for foremen, departmental heads, etc., are liable to be understated. The effect of this type of error can be partly eliminated by combining the data for employees in the more responsible occupations with those for their immediate inferiors.

The second type of error is due to unsound conclusions being drawn from the data rather than to actual mistakes in the data themselves. Some occupations demand a high standard of physical fitness and this constitutes a form of selection which operates not only in youth, when the initial choice of occupation will be influenced by a man's physique and general health, but also throughout working life, since a breakdown in health may necessitate

transfer to an occupation which does not demand so high a standard of fitness. Deaths among lives in the latter type of occupation may therefore be due to the occupational hazard of a calling which was abandoned before death. It is clear therefore that a mortality investigation on the normal lines may not give a true picture of the effect of occupation on mortality, although it may fairly portray the rates of mortality according to the occupation followed at the time of death. The general tendency is for the rates to be understated in heavy occupations such as mining in so far as these rates are regarded as indices of the occupational risks involved. This is one reason why the statistics published in the Decennial Supplement are not reliable guides for a life office to use for assessing extra premiums for occupation.

8. Occupation: female lives.

The occupational mortality of females was not investigated by the Registrar-General until the 1931 Decennial Supplement. Previously, it had been considered that there were far too many omissions and misstatements in the death returns to make such an investigation worth while, but in 1931 it was felt that there had been sufficient improvement in this respect to justify an investigation for spinsters, though not for married women.

As will be seen later, the mortality of married women classified according to their husbands' occupations was examined for a special purpose.

9. Social class or economic status.

From 1911 onwards, the data were divided into five classes based on a grouping of occupations which might be expected to include lives having roughly the same economic status. No. I comprised the upper and middle classes, No. III the skilled workers and No. V the unskilled workers; Nos. II and IV being intermediary. The grouping was not exactly the same on each occasion and some adjustments must be made before comparing the figures for the three investigations.

It is important to remember that, although the grouping into social classes is by occupation, the average social grade of the

employees in that occupation and not its relative position according to the mortality experienced determines the classification. The differences in the rates of mortality between one class and another are not therefore indicative of the effect of occupation on mortality.

10. Effect of the different factors on rates of mortality.

National statistics provide an excellent example of the difficulties of estimating the individual effects on mortality rates of a number of factors operating simultaneously (see Chapter XIII). At the 1931 investigation, the data were classified according to sex, subdivided by geographical situation and still further subdivided according to density of population. It was not practicable to make a further subdivision according to occupation, for there would then have been a very large number of classes and the data in an individual decennial age group would in many cases have been insufficient to produce reliable results. Furthermore, the work involved in making so many separate investigations would have been considerable.

It is outside the scope of this book to try to summarize all the conclusions which have been reached. From the standpoint of our subject, the conclusions themselves are less important than the reasoning which led up to them in face of the difficulties referred to above; the 1921 and 1931 Decennial Supplements, Parts I and II, are well worthy of study for this reason alone. A keen critical faculty is essential to a proper study of these publications, for the conclusions are necessarily of a tentative nature.

11. Use of control populations.

An interesting feature of the 1931 Supplement is the use of a *control population* in order to estimate the comparative effects of different occupations on mortality. A control population is one which is as nearly as possible similar to that under investigation, except that the particular factor whose effect we wish to estimate must not operate on the control population. The wives of men engaged in the occupations to be studied were chosen as a suitable control. In general, they should be affected by environment (i.e. density of population and housing conditions) and geographical

situation to the same extent as their husbands and, although sex has a marked effect on mortality, this effect is unlikely to vary much between one occupational group and another. It follows that differences between the mortality indices for men in a particular occupation and for their wives can reasonably be attributed to the effect of occupation, the indices being of course based on the appropriate standard rates of mortality for all men and all married women.

This method does not eliminate the element of physical selection referred to in para. 7 and consequently it does not always follow that the occupational factor has a more favourable or less adverse effect in the case of occupation A than in the case of occupation B, merely because the ratio of the mortality indices for men and for their wives is lower for A than for B.

It is, of course, quite possible to obtain a sound basis of comparison for some occupations without having to disentangle the effects of occupation on mortality from the effects of the other factors that we have discussed. A direct comparison is permissible if, for example, we are dealing with two or more occupations largely comprised in one industry, provided that the lives concerned largely belong to the same social grade and live in the same kind of district.

12. Conclusions drawn from the 1931 investigation.

Some of the principal conclusions advanced in the 1931 Supplement may be of interest, although, as already mentioned, such matters are really outside the scope of this book.

(*a*) "The immediate effects of occupation on men's mortality were of relatively slight importance compared with the environmental and economic conditions of home life." In other words, the fact that the lives in a particular occupation experience under- or over-average mortality may sometimes be due to favourable or adverse conditions inherent in the occupation, but more often it is due to the home conditions which are usually associated with the occupation, e.g. good or bad housing, residence in a closely or thinly populated area, high or low wages, etc. The distinction is important, for it determines the root cause which must be removed

if the mortality rates experienced in certain occupations are to be improved.

(*b*) "The worse the general environmental conditions the greater were the social contrasts in mortality for both sexes." These conditions tend to be worse in towns than in the country and it was noticeable that mortality varied to a far greater extent with social grade in the county boroughs than in the rural districts. In the rural districts some evidence of variation was observed among married women but not among men. This difference has been attributed to the fact that, while the adverse effects of poor housing and economic conditions in the lower social grades were neutralized for the men by the healthy nature of their occupations, their wives did not benefit to the same extent.

CHAPTER XXII

SICKNESS RATES

1. Definition of sickness.

It is not easy to give a simple definition of sickness in the sense in which the word is used when we talk of sickness rates. In order to understand its meaning, we require some knowledge of the nature of sickness benefits and the circumstances in which they are payable. Only a brief explanation can be given here and the reader is recommended to refer to chapters I and II of Brown and Taylor's *Friendly Societies* for a fuller description.

Sickness benefits are payable to members of friendly societies and other similar institutions who are unable to follow their normal occupations as a result of illness or accident. They usually take the form of weekly cash payments which continue as long as disability lasts, but the conditions of payment depend on the rules of the society. The meaning of "sickness" depends on the context in which the word is used, but in its narrowest sense it means disability which entitles the individual to sickness benefit from the particular society under review. This use of the word is apposite when we are investigating the sickness experience of a society, for we are not then concerned with periods of illness in respect of which no benefit is payable.

2. Essential difference between sickness rates and rates of decrement.

The rates with which we have been concerned in earlier chapters have all referred to decrements. In calculating the rates the units included in the numerators have disappeared from the experience or have been transferred to a different section, as for example, bachelors marrying in a combined marriage and mortality experience. Sickness rates, however, relate to a state rather than to an event, i.e. the length of sickness and not merely the event of falling sick. In a sickness rate the numerator consists of the total number of weeks sickness experienced between ages x and $x+1$

and the denominator is the number of years exposure to the risk of sickness contributed between these ages by the group of lives under review. Roughly speaking, therefore, the rate is the average number of weeks sickness experienced between ages x and $x+1$. The denominator is based on all the lives under review and not only those who experienced an attack of sickness between ages x and $x+1$.

It might be argued that a life falling sick should cease to contribute to the exposed to risk until he recovers. This procedure would, however, greatly increase the work of the investigation, for most attacks are comparatively short and a large proportion of the lives involved in an experience covering several years would have to be treated as entrants and exits more than once. Moreover, this procedure does not rest on a sounder theoretical basis than that in common use. Both methods are theoretically sound so long as the results are properly interpreted. In practice, therefore, the simpler procedure is invariably adopted.

Another distinction which should be noted in passing is that, whereas the fact of death, marriage or withdrawal can be definitely established, sickness can be feigned and the existence of a sound organization to examine claims is therefore an essential feature of a well run society.

3. Force of sickness and rates of sickness.

The *force of sickness* embodies a simpler idea than the force of mortality, for the latter cannot be measured until it is expanded to the equivalent annual rate, whereas the force of sickness at exact age x can be defined quite simply as the proportion of the lives attaining exact age x during the period of the investigation who were sick at that point of age. This function, which we shall denote by $\bar{z}_x$, is not of much value by itself; it is nevertheless the fundamental measure of sickness, just as μ_x is the fundamental measure of mortality.

The *rate of sickness* experienced by a group of lives aged x exactly may be defined as the ratio of the number of weeks sickness experienced by the group between exact ages x and $x+1$ to the

number of lives attaining exact age x, there being no causes of decrement other than death and no causes of increment.

The *central rate of sickness* may be defined in a similar manner, with the difference that the denominator is the number of years exposure between exact ages x and $x+1$ when each life is treated as exposed only up to the exact age at death.

If P_{x+t} represents the number of survivors to age $x+t$, where $0 \leqslant t \leqslant 1$, and the conditions laid down in the last two definitions are operating, the rate of sickness and the central rate of sickness, which are usually denoted by s_x and z_x respectively, may be expressed in terms of the force of sickness as follows:

$$s_x = \int_0^1 \frac{P_{x+t}\bar{z}_{x+t}\,dt}{P_x} \times 52\cdot18. \qquad \text{......(1)}$$

$$z_x = \frac{\int_0^1 P_{x+t}\bar{z}_{x+t}\,dt}{\int_0^1 P_{x+t}\,dt} \times 52\cdot18. \qquad \text{......(2)}$$

The rate is usually expressed in weeks and the multiplying factor 52·18 is the average number of weeks in the year, allowing for leap years.

Both rates are to some extent dependent on the rate of mortality to which the lives are subject, but the degree of the dependence is very small for the central rate. z_x is the weighted mean of the values of $\bar{z}_{x+t}$, multiplied by 52·18, the weights being the values of P_{x+t}, and a considerable variation in the rate of mortality with its resulting effect on the system of weights would be necessary to cause a significant change in the weighted mean, since the force of sickness will seldom vary much over a single year of age.

s_x depends on the rate of mortality to a greater extent for, unlike z_x, the denominator of s_x is independent of the rate of mortality at age x, so that there is no compensation when a change in mortality alters the numerator. The two rates are connected by the same relation as are q_x and m_x and we can express this approximately as follows:

$$s_x = z_x \times {}_{\frac{1}{2}}p_x,$$

a form which shows clearly the extent to which s_x depends on

mortality, since we can normally assume that z_x is independent of this factor.

4. Calculation of sickness rates.

In order to obtain the numerator of the ratio, the number of weeks sickness experienced by each life must be recorded according to the age at which it occurred. It is customary to use a six-day week, Sunday being a *dies non*, so that a period of sickness lasting from Friday till Tuesday week inclusive would count as 1 week 4 days. Further, it will be evident that one attack may extend over more than one year of age.

The denominator is usually obtained by an exposed to risk formula and the general method of procedure is the same as for a mortality investigation, but, as we shall see later, the individual terms in the formula may not be identical in the two cases.

The data in a sickness experience are nearly always grouped according to calendar year. In the first place, the work of allocating each attack of sickness to its exact year of age would be considerable, whereas it is a comparatively simple operation to allocate it to its correct calendar year and hence according to the assumed age at the beginning of that year. Secondly, the data will seldom include all the information necessary for an exact age method. A calendar year method is therefore generally adopted; the numerator of $\dot{s}_x$ then consists of the total number of weeks sickness experienced by lives in the calendar years at the beginning of which they were of assumed age x. The denominator is obtained by one of the methods of Chapter IV, suitably modified for the reasons given in the next paragraph.

5. Waiting period for new entrants: treatment of lapses.

In most societies there is a rule whereby a member is ineligible for sickness benefits in the first year or six months after entry. This is known as the *waiting period* and the date on which the member becomes eligible for sickness benefits is called the *date of freedom*. Death benefits may be available from the original date of entry or limited by a waiting period, possibly rather shorter than the waiting period for sickness benefits.

The date of entry or year of entry recorded by the society usually refers to original entry into membership. Suitable allowance must therefore be made in the exposed to risk formulae for the fact that exposure to the risk of sickness or death, as the case may be, did not begin until three, six or twelve months later. Failing this, we should include periods of exposure in the denominator for which there were no spells of sickness included in the numerator. Deaths within the appropriate waiting period should be treated as withdrawals.

The records should always be examined carefully to make sure that it is the date or year of original entry which is given and not the date or year of freedom. If the latter is given, this date or year can be used without adjustment in the same way as the date or year of entry in a mortality experience of assured lives.

When a member's contributions fall into arrear, he may immediately cease to be eligible for sickness benefits, although his membership will not lapse finally until contributions have been in arrear for six months or a year, during which time he will be entitled to reinstatement, possibly subject to evidence of health. The right to death benefits may continue for three or six months after contributions cease or may even be allowed right up to the time when the member's name is removed from the books. Once again, it is important for the actuary to find out what is treated as the date or year of withdrawal in the society's records.

6. Exposed to risk formulae.

In order to show how allowance is made for the factors referred to in the last paragraph, we shall develop an exposed to risk formula. Suppose that (*a*) the waiting period is six months for sickness, (*b*) there is no waiting period for deaths, (*c*) eligibility for sickness benefits ceases as soon as contributions are outstanding, (*d*) the society covers death benefits for three months longer, (*e*) membership ceases after a further six months.

Let us assume that we are given the age next birthday at entry and the calendar years of entry and exit. We shall take year of entry to mean the year of original entry and year of withdrawal to be the

year in which the member ceases to be eligible for death benefits. We then group together at assumed age x all movements of a particular type, e.g. beginners, deaths, etc., where x is defined as follows:

Age next birthday at entry $-1+$ calendar year of movement $-$ calendar year of entry.

On the average, x will be the exact age at the beginning of the calendar year of the movement. (It should be noted, however, that for enders, assuming that the investigation ends on 31st December 1935, we must treat the year of exit as 1936, in which case x will be the average age on 1st January 1936.)

In the same way, we sum the total number of weeks sickness at assumed age x such that x is:

Age next birthday at entry $-1+$ calendar year of sickness $-$ calendar year of entry.

The new entrants n_x enter at age $x+1$ next birthday and the assumed age at the beginning of the calendar year of entry is x. Assuming that dates of entry are evenly spread over the calendar year, n_x are exposed to the risk of death for half a year at assumed age x and the appropriate term in ${}^dE_x$ (the exposed to risk of death) is $\frac{1}{2}(n_{x-1}+n_x)$.

Exposure to the risk of sickness does not begin until six months later. For lives entering in calendar year k, the dates of freedom lie between 1st July in year k and 30th June in year $k+1$, the average date being 1st January in year $k+1$ if dates of entry are uniformly distributed. On the average, therefore, the new entrants at age x next birthday become eligible for sickness benefits on the 1st January following entry, on which date their assumed age is x by definition.

In the case of withdrawals, the term in ${}^dE_x$ is $-\frac{1}{2}(w_{x-1}+w_x)$, since w_x are on the average exposed to the risk of death for half a year at assumed age x. These lives ceased to be exposed to the risk of sickness three months earlier and it is usual to treat them as exposed for three months on the average at assumed age x so that the term for withdrawals in ${}^sE_x$ is $-\frac{1}{4}(w_{x-1}+3w_x)$.

. The exposed to risk formulae are, then,

$$^dE_x = {^dE_{x-1}} + b_x - e_x + \tfrac{1}{2}(n_{x-1} + n_x) - \tfrac{1}{2}(w_{x-1} + w_x) - \theta_{x-1},$$

$$^sE_x = {^sE_{x-1}} + b_x - e_x + n_{x-1} - \tfrac{1}{4}(w_{x-1} + 3w_x) - \theta_{x-1},$$

$$^sE^c_x = {^sE^c_{x-1}} + b_x - e_x + n_{x-1} - \tfrac{1}{4}(w_{x-1} + 3w_x) - \tfrac{1}{2}(\theta_{x-1} + \theta_x).$$

If we also require rates of withdrawal, the corresponding formula will be as follows, assuming once again that a member is treated as withdrawing when he ceases to be eligible for any benefits, i.e. three months after contributions fall into arrear.

$$^wE_x = {^wE_{x-1}} + b_x - e_x + \tfrac{1}{2}(n_{x-1} + n_x) - \tfrac{1}{2}(\theta_{x-1} + \theta_x) - w_{x-1}.$$

It should be noted that the formula for ^sE_x involves an error in respect of the new entrants in the calendar year immediately before the beginning of the experience. All these lives, provided that they were still members, would be included as beginners in ^sE_x as well as in ^dE_x, although those who entered in the second half of the year would not reach their dates of freedom until after the beginning of the investigation period. ^sE_x is accordingly overstated. No exposure is however included in ^sE_x for new entrants in the last year of the experience (they will be included both as new entrants and as enders at the same age), although those who entered in the first half of that year would reach their dates of freedom before the investigation ended. It is usual to assume that these errors will offset each other and they are accordingly neglected.

Members who have been temporarily ineligible for benefits owing to non-payment of contributions but who were reinstated within the statutory period—in our example, nine months after contributions ceased—are normally included in the exposed to risk as if there had not been any temporary loss of cover. The rates are thus understated, but the error will usually be small. Moreover, from the financial standpoint, even a fairly large error will not cause any loss to the society if the rates are used to calculate future rates of contribution, unless there is in future a marked reduction in the number or duration of the periods off the risk. The practice of neglecting these periods is consistent with the practice followed in life office investigations for policies which have lapsed but which were revived during the investigation period for their full sums assured.

No new principles are involved in the methods of constructing exposed to risk formulae for the calculation of sickness and mortality rates from the experience of friendly societies. It is therefore unnecessary to give any more examples of these methods, as the reader should have no serious difficulty in devising suitable formulae, provided that the basic principles are kept in view.

7. Periods of sickness pay.

When an attack of sickness has lasted for a certain period, e.g. six months, the rate of sickness benefit—or *sickness pay* as it is often called—is usually reduced. This practice not only provides some protection against serious loss to the society from an excessively high proportion of claims of long duration but also discourages malingering. The actual rates of pay which are granted vary considerably between different societies, but a typical scale is: 20*s*. per week for the first six months of any one attack and 10*s*. per week during the remainder of the attack. In some societies, there are more than two rates of pay, e.g. 20*s*. per week for the first six months, 15*s*. for the second six months and 10*s*. thereafter.

The lowest rate of benefit, which usually continues as long as sickness lasts, is sometimes called *continuous sickness benefit* but various other names are used.

8. Off period.

Failing some rule to prevent it, there would be a risk of members in receipt of continuous benefit signing off before they had recovered and going sick again after a week or two, thus securing the highest rate of pay instead of the lowest. It is, therefore, customary to provide for the linking together of two attacks separated by less than a fixed period and treating them as a single attack for the purpose of determining the appropriate rate of pay. This fixed period is called the *off period* and its usual duration is one year, although shorter or longer off periods are not uncommon.

To illustrate its operation, consider the following series of attacks, the off period being one year and the rate of pay 20*s*. for

the first three months, 15*s*. for the next three months and 10*s*. thereafter.

No.	Limiting dates	Duration of sickness					
		1st period		2nd period		3rd period	
		W.	D.	W.	D.	W.	D.
1	1st Feb. 1938 to 15th March 1938	6	1	—	—	—	—
2	12th April 1939 to 19th Aug. 1939	13	—	5	4	—	—
3	3rd Jan. 1940 to 20th May 1940	—	—	7	2	12	3
4	12th Sept. 1940 to 5th Oct. 1940	—	—	—	—	3	3
5	6th Oct. 1941 to 4th Dec. 1941	8	4	—	—	—	—

More than a year separated the first two attacks and the member was therefore eligible for the highest rate when he fell sick in April 1939. The third attack was however linked with the second and the fourth with the third, since both the intervening periods were less than a year. These three attacks, therefore, ranked as one attack for the purpose of determining the rates of pay. The last attack ranked as a new attack, however, and in this connection it is noteworthy that the intervening period was only one day over a year. This suggests that the member may have fallen ill shortly before 6th October 1941 but deferred making a claim until he had qualified for full pay—a proceeding which was probably within the letter although not in accordance with the spirit of the rules.

9. Subdivision of sickness rates according to period of pay.

Rates of sickness are normally calculated for each separate period of pay. The numerator of each ratio consists of the number of weeks sickness at the appropriate rate of pay, but the denominator is the same in each case, i.e. the number of years exposure to the risk of sickness regardless of the rate of pay for which the member was eligible at the time. Hence the sum of the rates for the different periods gives the rate for "all sickness".

It may be argued that it is unsound to include a member in the exposed to risk for first period sickness (i.e. sickness which entitles

the member to full pay) at a time when he is ineligible for full pay, i.e. during an off period or during that part of an attack which oversteps the first period. The alternative would involve treating each member as a withdrawal on his ceasing to be eligible for full pay and as a new entrant at the end of the off period in calculating the exposed to risk of first period sickness. On withdrawal from the exposed to risk of first period sickness, the member would be treated as a new entrant for second period sickness. The exposed to risk for the various periods would add up to that for all sickness, but the rates would not be additive. Rates of transfer between one grade and another would have to be calculated as in the case of a graded service table and the whole process would be far too complicated to be practicable.

Special symbols are used to denote rates of sickness for different periods. If n is the length in weeks of the period in question and m the number of weeks which an attack must last before the particular period is reached, the rate is denoted by $s_x^{m/n}$. For example, $s_x^{26/26}$ would normally be used to represent the rate of second period sickness, if the rate of pay was reduced after an attack had lasted six months and again after it had lasted a further six months. $s_x^{52/\text{all}}$ would be used to denote the rate of continuous sickness, if the lowest rate of pay were reached after an attack had lasted for a year.

It is most important that the student should realize that the subdivision of sickness rates according to period does not imply any inherent distinction in the types of sickness applicable to the different periods. The lengths of the periods depend on the rules of the society. Even if the periods are the same for two societies, the rates may not be comparable owing to a difference in the off period with the result that an attack falling into the first period in one society would have been allocated to the second period if the rules of the other society had been in force.

10. Collection and tabulation of data.

Individual cards are generally used for collecting the data and the basic details are usually inserted by the officials of the society, especially when the experience of a group of affiliated societies is under investigation. The societies cannot be expected to exercise

the degree of skill and care in writing the cards which one would be entitled to expect in a life office investigation. The form of the card should therefore be as simple as possible and the instructions for filling up the cards should be couched in clear and unambiguous language. As a rule, it is better that societies should insert only those particulars which are obtainable direct from their books. It would be left to the actuary to calculate the assumed ages of entering and leaving the experience.

Particulars of the sickness experienced by each member are, of course, required for each calendar year. It is an advantage if the society records the number of weeks and days at each rate of pay so that these particulars may be transcribed on to the cards by the officials of the society. Failing this, the total amount paid in sickness benefits in each year at each rate of pay may be recorded for each member and this information should then be recorded on the cards instead of the duration of sickness; the latter function will be obtainable at the most convenient stage by dividing by the rate of pay.

In the unlikely event of neither of these figures being directly obtainable, it may be desirable to instruct the societies to insert merely the dates when each attack began and ended. The actuary will then be responsible for the tedious work of calculating the number of weeks sickness and allocating it to its correct period. For this purpose, details of the sickness experience of each member for a year or two before the investigation will be required, so that allowance can be made for the effect of the off period in linking up two attacks, the first of which ended before the investigation began. Fortunately, it is seldom necessary to have recourse to the actual dates of sickness attacks, except when an investigation is required into the experience of a group of societies having different periods of sickness pay or different off periods (see para. 17).

The form of card will depend on the data available, but the following example shows a typical layout.

Exact dates of birth, entry and exit are assumed to be unknown. As the year of birth is known, method III of Chapter IV could be employed and it would not be necessary to record the age next birthday at entry on the card. This information might be required

for other purposes, however, and in any case might be used as a rough check on the correctness of the recorded years of birth and entry. The box headed "Age" on the front and the columns headed "Age" on the back would be filled in by the actuary.

Front

No. .... Name ....

Occupation .. . . Sex ....

	Age
Year of birth	
Year of entry	. ...
Year of exit	.. ..

Cause of exit ....

Age n.b.d. at entry ....

Remarks

Back

		Sickness claims						
		Period						
Year	Age	1st		2nd		3rd		
		W.	D.	W.	D.	W.	D.	
1930								
1931								
1932								
1933								
1934								
Total								

To obtain the exposed to risk, the cards would be sorted and subsorted and the numbers tabulated in the same way as for a mortality investigation. A schedule would be prepared for each age and particulars of the sickness which had been recorded on the backs of the cards would be entered in the appropriate schedules, from which the total sickness at each age would then be obtainable for each period of pay and for all periods combined. The required rates would then be calculated.

In a large investigation, use could be made of mechanical devices to speed up the work of sorting and tabulating.

Cards are sometimes kept by a society for recording the information needed for valuation purposes and these may be suitable for a mortality investigation. The same card would be used throughout the whole duration of membership and, in order to avoid entering the age each time a valuation is made, the valuation year of birth,

which remains constant throughout life, would be recorded; at any valuation date the lives having the same valuation date of birth would also be of the same attained age. The same applies when the cards are used for grouping movements at the same age in a sickness investigation.

11. Factors influencing sickness rates.

As in the case of mortality rates, a number of factors such as sex, geographical situation, density of population and occupation have an effect on sickness rates, but the relative importance of the factors is not the same in the two cases.

(*a*) *Sex*. Sickness rates for women tend to be higher than for men and those for married women higher than for spinsters. The experience of men and women should always be investigated separately, especially as the rates of benefit will usually differ. The segregation of single and married women would necessitate the calculation of marriage rates before the results could be used either for the calculation of rates of contribution or for valuation purposes and it is therefore usual to neglect marital status. From the financial standpoint this procedure is satisfactory unless there is reason to expect that in future the proportion of married women age by age will differ considerably from the proportions existing among the data used in the investigation.

In this connection, it should be remembered that, so far as sickness rates are concerned, marriage is not merely a selective influence but has a direct bearing on the experience. As mentioned previously, the normal qualification for sickness benefit is the inability of the individual to follow his or her normal occupation. The chief protection against malingering lies in the fact that the amount of benefit is usually substantially less than the wages lost by absence from work. This factor loses some of its force in the case of a married woman whose husband is employed and is doubtless the principal reason why the sickness experience of married women tends to be unfavourable.

(*b*) *Geographical situation, density of population and occupation.* In the Manchester Unity experience 1893–97, Watson made a careful investigation of the effect of these factors. The country was

divided into three sections based on an industrial rather than a purely geographical grouping—that is to say counties whose populations were primarily engaged in the same types of industry were placed in the same section. Each section was subdivided into rural and urban subsections, i.e. according to density of population, and the lives for each subsection were further split up into nine occupational groups.

By examining various combinations of the data, conclusions were reached as to the relative importance of the different factors in their effects on sickness and mortality rates. Watson decided that, whereas geographical situation and density of population appeared to have a significant effect on mortality rates, their effect on sickness rates was comparatively slight. It appeared however that occupation had a marked effect on sickness rates but seemed to be less important than the other factors in its influence on mortality (cf. the conclusions derived from national statistics, Chapter XXI).

Eventually, Watson calculated sickness rates for four occupational groups, obtained by combining several of the original groups, and then calculated monetary functions by grafting each set of rates on to each of three sets of mortality tables based on the three geographical groups, rural and urban districts combined. The reader is recommended to study the official account of the experience describing in detail the reasons for the combinations adopted.

At first sight the combination of sickness rates based on one set of lives with mortality rates based on a different set may appear to be unsound. The object was, however, not so much to obtain tables applicable to particular groups of lives as to set up standard tables based on substantially different levels of sickness and mortality rates, so that each society would in future be able to choose a standard monetary table, corresponding roughly to its own experience, which it could use for calculating contributions and for valuation purposes.

In the case of a small society or an individual branch of a large society, it will seldom be necessary to make many subdivisions of the data when investigating the sickness and mortality experience. All the members will usually live in the same district and in some

cases the majority will be engaged in the same or similar occupations. Heterogeneity does not therefore present any serious problems except in an investigation of the experience of a whole affiliated order or of a society having branches in different localities.

12. Advantage of central sickness rates.

When there is a possibility that the sickness rates from the experience under review will be combined with the mortality rates obtained from a different body of lives, it is advisable to calculate central sickness rates, since these rates are virtually independent of mortality (see para. 3). The value of a sickness benefit of a unit per week at age $x+t$ to a life now aged x will then be $v^{t+\frac{1}{2}} L'_{x+t} z_{x+t}/l'_x$, where the accented functions relate to the mortality table with which the sickness rates are to be combined.

If ordinary sickness rates were used, the corresponding expression would be $(v^{t+\frac{1}{2}} l'_{x+t} s_{x+t}/l'_x)\,({}_{\frac{1}{2}}p'_{x+t}/{}_{\frac{1}{2}}p_{x+t})$, the last term being necessary to correct for the fact that s_{x+t} is dependent on the rate of mortality experienced by the lives on which the sickness rates are based. Clearly, this expression is less convenient for purposes of calculation than that based on the central rate.

13. Proportion sick in a year.

This function is sometimes calculated in addition to the rate of sickness. As applied to a particular period of pay, it is the ratio of the number of lives receiving sickness pay at that rate at any age between x and $x+1$ to the average number of lives between those ages. In calculating the function, the denominator is the same as the central exposed to risk of sickness. When the proportion sick between ages x and $x+1$ for all periods of sickness combined is calculated, no life is included more than once in the numerator at a particular age, even although sickness pay at more than one rate may have been received at that age. The sum of the proportions sick at the various periods therefore exceeds the proportion for all sickness.

It should be noted that, if a single attack extends over two successive ages, it will affect the proportion sick at both ages.

The ratio of the central rate of sickness at age x to the proportion sick at age x gives the average number of weeks sickness at age x

experienced by those who claimed the particular rate of sickness pay. This ratio provides a rough measure of the average duration of sickness attacks at that age for the appropriate period and is useful in tracing the reason for an increase in the amount paid in sickness claims, i.e. whether it is due to more or longer attacks of sickness or both. It is not an exact measure of the average length of a complete attack, as some claims will last for more than a year.

Comparison of the proportions sick in the various periods with the corresponding proportions in the past may also indicate whether the duration of the attacks has been increasing.

The proportion sick in a year should not be confused with the force of sickness defined in para. 3.

14. Comparison of actual and expected sickness.

Such a comparison may be carried out in the same way as the comparison between actual and expected deaths. When investigating the experience of any society except a very large one, the data will be insufficient to justify the construction of a sickness table and a comparison of the actual sickness with the expected by one or more standard tables will indicate what standard table should be used—with adjustments if necessary—to calculate contributions, etc. As a rule, the standard table employed for comparison will be that used in the past for calculating monetary functions.

A comparison on these lines made each year provides an indication of the trend of the sickness experience from one year to another and gives an early warning of a deterioration in the rates, such as might be caused by a relaxation of the precautions against malingering.

In order to test the financial effect of the experience, cost of sickness may be compared instead of rates of sickness. If, for example, the rates experienced for the period of full pay were unfavourable while those for subsequent periods were favourable, the total cost might be greater than the expected although the total amounts of actual and expected sickness were roughly the same.

15. Selection.

Temporary initial selection has only a small influence on sickness rates. Friendly societies do not demand from new entrants evidence

of health such as would be required by a life office from its proposers for new assurances; in general, therefore, temporary initial selection exists only from the fact that all lives are free from sickness at the time of entry. Sickness tables are therefore invariably constructed on an aggregate basis.

Duration since entry has some effect on the distribution of sickness rates over the periods of pay. The most obvious example is that of a group of lives who have not been members for long enough to have qualified for the lower rates of pay. If full pay lasts for the first six months of an attack, half pay for the second six months and quarter pay for the remainder, a group of lives within six months of their dates of freedom would not experience any sickness except at full pay, while a group of lives within a year of their dates of freedom would not experience any quarter pay sickness.

If the lives had entered sooner, some of the sickness which attracted full pay might have been linked up with previous sickness as a result of which it would have been allocated wholly or partly to a later period of pay. It will be seen, therefore, that this form of selection is purely statistical in origin and is not the result of any inherent difference in the type of sickness experienced by lives at the short durations. It is in fact entirely due to the definition of sickness, which determines the subdivision into periods in a purely artificial way. This feature must be borne in mind if we are investigating the experience of a very young society or one which has recently had an exceptionally heavy influx of new members at any other than the youngest ages.

Similarly, if we examine the sickness experience at the old ages, we may find what would at first sight seem to be a surprisingly low rate for first period sickness. This phenomenon is due to the fact that a proportion of the members are in permanent receipt of sickness benefit at the lowest rate and are therefore ineligible for benefit at the higher rates.

16. Special types of sickness benefit.

Some societies incorporate in their rules special provisions to restrict the loss incurred from long attacks of sickness. The following are examples of the more common provisions of this kind.

(*a*) Restriction of continuous sickness benefit to a maximum period of, say, six months at a time. At the end of this period, sickness benefits will cease, but after a further year the member will again become eligible for pay at the full rate. The year during which no pay is allowed must not be confused with the off period. A member who is permanently disabled receives benefit in a recurring cycle, e.g. full pay for six months, half pay for six months, no pay for a year, full pay for six months, etc.

(*b*) Full pay continuing until the total benefits received equal the total amount paid in contributions, after which a lower rate of pay is allowed. A member who is eligible for half pay only will not qualify for full pay again until his total contributions once more exceed the total benefits he has received.

If the sole object of the investigation is to calculate sickness rates according to the definition of sickness applicable to the particular society, no difficulty should arise. Sickness will be allocated to the first and second period according to the rate of pay which was received and, in the case of (*a*), sickness which did not qualify for pay will not be included at all.

If, on the other hand, we wish to compare the experience with a standard table based on more normal rules, the procedure is not so straightforward and much will depend on the nature of the differences between the two sets of rules. Suppose, for example, that rates according to the standard table are divided into the following periods: first six months, second six months, second year and over two years, with an off period of one year, while the rules of the society under review are in accordance with example (*a*) above, i.e. full pay for six months, half pay for six months, no pay for a year, also with an off period of one year. If full details were obtainable of the sickness which did not attract any benefit, it would be possible to redistribute the sickness into the periods of pay according to the standard table and a direct comparison would then be feasible.

This comparison would not however be entirely satisfactory. It would show how the society's experience would have compared with the standard table, if the rules on which the latter table was based had been in force. What we really require however is a comparison of the sickness rates according to the rules of the society

whose experience we are investigating. For this purpose, it is the standard rates which should be brought into line with the experience rates and not the converse. To do this accurately, we should have to examine the crude data on which the standard table was based, so as to ascertain what benefits would have been paid if the rules of the society had been in force. A new standard table would, in fact, have to be constructed; this is of course out of the question.

Nevertheless, it is sometimes possible to obtain a fairly satisfactory comparison of the total actual cost of all sickness benefits with the expected cost. In the above example, the expected cost might be taken as one unit times the rate for "first six months" sickness plus half a unit times the rate for "second six months" sickness plus three-eighths of a unit times the rate for "over eighteen months" sickness. Three-eighths of a unit is assumed to be the average rate of pay for continuous sickness, since over a complete two years cycle only six months full pay and six months half pay would be received.

The reason why the "over eighteen months" sickness is chosen as one of the comparative measures should be noted. No benefit is payable by the society during the second, fourth, etc., years of sickness and the benefit payable during the third, fifth, etc. years is treated for comparative purposes as if it were averaged. We might treat the benefit paid in the third year as if it were averaged over the second and third years, in which case the correct period of sickness for comparison would be "over one year". Alternatively, we might treat the third year's benefit as if it were averaged over the third and fourth years, in which case comparison should be made with "over two years" sickness. It must be remembered, however, that the attacks of sickness will not in general last for a period which is an exact multiple of two years and it is therefore necessary for comparative purposes to spread the actual payments in such a way that the weighting will not be seriously disturbed. Hence the third year's benefit is spread over the period from the end of one and a half to the end of three and a half years and a more nearly correct period of sickness for comparison is "over eighteen months". The rates for this period are not obtainable direct from the standard table, but could be found with sufficient accuracy by interpolation.

Most of the problems which arise in connection with special sickness benefits concern rates of contribution and valuation and we need not consider such matters here. It is just as well that the reader should understand however that even when a comparison of the total actual cost of all sickness with the expected is possible, this is not always a sufficient test to determine the basis for monetary functions. The rules of the society may be such that the experience of the past is not a reliable guide to the future. In example (*b*), for instance, a steady improvement in the sickness rates at the younger ages during the past twenty years might lead to an increase in the cost of full pay and a fall in the cost of half pay at the older ages in future, since a larger proportion of lives would be eligible for full pay at the latter ages.

17. Standard tables.

Several references have already been made to the Manchester Unity investigation of 1893–97 and this is the only experience which the reader is recommended to study. Incidentally, some information about earlier investigations is given in the official account. The peculiar value of the M.U. tables arises from the nature of the organization, which comprises a large number of lodges which no doubt experience different rates of mortality and sickness. A series of standard tables was accordingly required, instead of a single table, so that each lodge would be able to find a table suitable for its own use with the minimum of adjustment. The tables which were prepared have proved a great boon to many friendly societies outside as well as inside the Manchester Unity.

The individual societies had their own rules and in order to obtain the combined experience full details of all sickness had to be collected so that each attack could be divided into the appropriate periods of sickness according to a standard set of rules. If the experience for full pay, half pay, etc. had been aggregated irrespective of the period of such pay and the length of the off period, the resulting rates, other than those for all sickness, would have been meaningless.

The methods used in calculating the rates are clearly explained in the official account and, although the exposed to risk formulae

are not given, the reader should have no difficulty in working them out. If the formulae are applied to the published data, some minor discrepancies will be found in the exposed to risk at the young ages.

18. Analysis of sickness rates into short periods.

For purposes of the National Health Insurance scheme, sickness rates subdivided into short periods of attack were required, i.e. first three days, second three days, second week and so on. As no existing table was subdivided into such short periods, Manchester Unity rates were split up for the purpose by a process described in Cmd. 6907. This process is complicated and some of the steps taken are not clearly described in the memorandum. The following is a résumé of what appears to be the rationale of the process.

Let y_t be the proportion of lives between ages x and $x+1$ who experienced sickness of exact duration t between those ages. The denominator of y_t is the average number of lives exposed to the risk of any sickness between ages x and $x+1$, i.e. the same as the denominator of z_x, the central rate of sickness at age x, and also the denominator of the central rate of sickness for the various periods of attack. y_t must not be confused with $\bar{z}_{x+t}$ which is the ratio of the number sick at exact age $x+t$ to the number exposed to risk at exact age $x+t$. In $\bar{z}_{x+t}$, t refers to age whereas in y_t, t is the duration since the beginning of the attack.

The expression $\int_t^\infty y_t\,dt$ gives the value of s_t by which we shall denote the central rate of sickness at age x for all durations in excess of t. This follows from the fact that $y_t \Delta t$ is, by the definition of y_t, the rate of sickness between durations t and $t+\Delta t$, Δt being a small quantity. If t is expressed in weeks, $\int_{104}^\infty y_t\,dt$ is the rate for "over two years" sickness and so on.

If now we can construct a graph of y_t, the sickness rate for any given period of attack will be obtainable by measuring the area enclosed by the axis of t, the curve of y_t and the appropriate ordinates. Thus the area bounded by the ordinates $t=26$ and $t=52$ would be the rate for the second six months of sickness. Before we can draw the graph, however, we must obtain a sufficient number of values of y_t.

The most important individual value is y_0, the proportion of lives entering on a new attack of sickness between ages x and $x+1$. y_0 is the difference between the proportion of lives experiencing first period sickness at any age between x and $x+1$ (as defined in para. 13) and the proportion experiencing first period sickness at exact age x divided by ${}_{\frac{1}{2}}p_x$ (to make the denominator the same as in the other two functions). The first function is obtainable direct from the M.U. tables and a good approximation to $\bar{z}_x^{13}$ in the second is $\dfrac{z_{x-\frac{1}{2}}^{13}}{52\cdot 18}$. The rationale of this approximation is apparent from relation (2) of para. 3. It will be seen that an error in the value of y_0 is introduced by the fact that some of the lives contributing to the numerator of the proportion experiencing first period sickness at any age between x and $x+1$ were not sick at exact age x, but experienced sickness at age $x-1$ with which the attack at age x was linked by the off period. Hence these attacks at age x were not new attacks which should contribute to y_0. The error is not however likely to be serious.

By means of the identity $D \equiv \log(1+\Delta)$, the expression $y_t = -\dfrac{ds_t}{dt}$ can be expanded in terms of Δs_t (the ordinary first, second, etc. differences with regard to t) and we therefore require values of s_t at unit intervals in order to obtain values of Δs_t. s_t is obtainable from the M.U. tables for $t=1, 2, 4, 8$, the unit being thirteen weeks, and we must interpolate for $t=3, 5, 6, 7$. This can be done by the use of divided differences, but it is simpler to change the variable to n where $t=2^n$. The values of n corresponding to the values 1, 2, 4, 8 of t are 0, 1, 2, 3 and a simple interpolation can be made with regard to n to obtain values for $n=\dfrac{\log 3}{\log 2}, \dfrac{\log 5}{\log 2}$, etc. (corresponding to $t=3, 5$, etc.). Actually y_t was expressed in the equivalent form $-s_t\dfrac{d \log s_t}{dt}$ so that the interpolation could be performed on $\log s_t$ instead of s_t, the advantage being that the first differences of $\log s_t$ are more nearly in a straight line than those of s_t. Having obtained s_t for all integral values of t from 0 to 8, the differences can easily be extracted. As it appeared that $\Delta \log s$

tended to a constant for each age group, constant values were assigned to $\Delta \log s_t$ for values of t in excess of 7. The resulting values of the various differences can be applied to find the values of y_t for integral values of t from 0 to 8 from the formula

$$y_t = -s_t \log (1+\Delta) \log s_t$$
$$= -s_t\left(\Delta - \frac{\Delta^2}{2} + \frac{\Delta^3}{3} - \ldots\right) \log s_t.$$

As the unit is 13 weeks, the values of s_t must be expressed in terms of this unit before they are used in the formula, but for $\log s_t$ this is unnecessary since we are only concerned with first and higher differences, and clearly $\Delta \log ks_t = \Delta \log s_t$ when k is constant.

The nine values of y_t are sufficient to permit of the graph being drawn and the sickness rate for the period between any two durations of attack up to two years can be obtained by estimating the appropriate area. Here again, it must be remembered that the unit is 13 weeks, so that the area between the ordinates $t=\frac{3}{13}$ and $t=\frac{4}{13}$ gives the rate for fourth week sickness. This area is approximately $\frac{1}{13}\{\frac{1}{2}(y_{\frac{3}{13}} + y_{\frac{4}{13}})\}$. As mentioned above, the unit in which this function is expressed is 13 weeks and if, in accordance with normal practice, the rate is expressed in weeks, the function takes the simple form $\frac{1}{2}(y_{\frac{3}{13}} + y_{\frac{4}{13}})$.

The student is recommended not to try to remember the details of the process. They are quite unimportant and have been dealt with at some length here merely in order to satisfy anyone who has been puzzled by Cmd. 6907. The underlying principles are important, however, and the reader should make sure that he understands exactly what y_t and s_t represent and why they are connected by the relation $y_t = -\frac{ds_t}{dt}$. In particular, the exact meaning of y_0 and the method by which it was estimated should be clearly appreciated. Treated in this way, the memorandum provides a useful test of the student's understanding of the various functions introduced in this chapter.